Registration of Tomographic Images to X-ray Projections for Use in Image Guided Interventions

Penney, Graeme

Awarding institution:
King's College London

Title: Registration of Tomographic Images to X-ray Projections for Use in Image Guided Interventions

Author: Graeme Penney

Registration of Tomographic Images to X-ray Projections for Use in Image Guided Interventions

Graeme Patrick Penney

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of London

December 1999

Computational Imaging Science Group,

Division of Radiological Sciences and Medical Engineering,

Guy's, King's and St. Thomas' School of Medicine,
King's College London

Abstract

This thesis contributes to knowledge by describing a new method to allow information from a pre-operative 3D modality to be used during an intervention which is guided using 2D fluoroscopy images.

An algorithm has been designed to obtain the pose of a CT volume with respect to a single fluoroscopy image. The registration algorithm is based on the production of digitally reconstructed radiographs from the CT volume, which are compared to the fluoroscopy image using a similarity measure. The novelty of the work described in this thesis is in both the design of the registration algorithm and also in investigating the specific requirements placed upon a similarity measure when attempting to register a pre-operative CT volume to an interventional fluoroscopy image. Seven similarity measures were investigated. Experiments were carried out to calculate the accuracy and robustness of the registration algorithm using each similarity measure. Initially fluoroscopy and CT images of a lumbar spine phantom were used. The accuracy of the registration algorithm was calculated by comparing the final registration positions with a "gold-standard" registration calculated using fiducial markers. More realistic datasets were simulated using the phantom fluoroscopy image with clinical image features overlaid. Results show that the introduction of soft-tissue structures and interventional instruments into the phantom image can have a large effect on the performance of some similarity measures previously applied to 2D-3D image registration. The similarity measures were also tested on clinical data from aortic stenting procedures, where *k-fold* cross validation was used to obtain an estimate of the registration accuracy. The results from these experiments showed that two measures were able to register accurately (RMS rotational error of 0.76° and RMS in-plane translational error of 0.85mm) and robustly (10% failure rate) even when soft-tissue structures and interventional instruments were present as differences between the images. These two measures were pattern intensity and gradient difference.

Finally the thesis describes a novel combination of the 2D-3D registration algorithm with a deformation algorithm. The registration algorithm was used to obtain information on the relative movement of the vertebrae between the pre-operative CT image and interventional fluoroscopy image. This information was then used to warp the pre-operative modality so that it more accurately represented the intra-operative scene.

Acknowledgements

Firstly I would like to thank my supervisors, David Hawkes and Derek Hill, for the initial inspiration and all their help and encouragement. A big thank you also goes to my unofficial supervisor, John Little, who helped me a great deal while he was working at Guy's and still puts up with my questions now he has left. I am also very grateful to the rest of the Computational Imaging Science Group at Guy's Hospital, past and present. Special thanks goes to Eddie Edwards, Colin Studholme, Daniel Rückert, Colin Renshaw, Ron Gaston, Steve Keevil, Alan Munday[1], Matt Clarkson, Andy Castellano Smith, Jane Blackall, Andy King, Philipp Batchelor, Dave Atkinson and (last but certainly not least) Calvin Maurer Jr — whose tree trunk legs will always be remembered playing for "Where's Junior". Thank you all; for your patience in answering my questions and for asking me, on occasion, "what are you having?".

I am very grateful to the clinicians who have shared their expertise and valuable time with me, particular thanks go to Andy Adam and Mark Cowling. I would also like to thank all the radiographers at Guy's Hospital who have provided invaluable help in acquiring the images I required for my work.

I was sponsored by Philips Medical System EasyVision Advanced Development during the course of my PhD: thank you, firstly for the funding but also for the helpful advice and software, particular thanks go to Frans Gerritsen, Hans Burrman, Karel Strasters, Kees Visser, Pol Desmedt and Marcel Quist. I would also like to thank Jürgen Weese from Philips Research Hamburg for what I believe (and I hope he does too) has been a very useful collaboration

Thanks to Mum and Dad for everything they have done for me — the page isn't big enough for that one. A big thank you also goes to Robin for putting up with me for the last three (plus) years and hopefully will continue doing so for a few more years.

[1] social member

Contents

List of Tables

List of Figures

List of Abbreviations

2D	Two Dimensional
3D	Three Dimensional
CT	(X-ray) Computed Tomography
AP	Anterior-Posterior
DRR	Digitally Reconstructed Radiograph
DSA	Digital Subtraction Angiography
MR	Magnetic Resonance
PA	Posterior-Anterior
PET	Positron Emission Tomography
PLN	Percutaneous Laser Nucleotomy
RF	Radio Frequency
RMS	Root Mean Square
ROI	Region Of Interest
SVD	Singular Value Decomposition
TLE_o	Target Localisation Error
TRE_o	observed Target Registration Error
TRE_t	true Target Registration Error

Chapter 1

Introduction and summary

The work described in this thesis is concerned with the registration of a three-dimensional (3D) tomographic image with a two-dimensional (2D) x-ray projection image. The aim is to register a pre-operative computed tomography (CT) image with an intra-operative fluoroscopy image. This registration will enable information from the pre-operative image to directly aid needle or stent placement during an intervention carried out using fluoroscopic guidance.

Registering a tomographic image to an x-ray projection image is often referred to as a 2D-3D image registration. This chapter initially provides a definition of 2D-3D image registration. The next section states the major aim of the work described in this thesis and discusses the clinical motivation behind this aim. The last section in this chapter provides an overall summary of the thesis.

1.1 2D-3D image registration

Registration is defined in Maurer *et al.* [88] as *"the determination of a one-to-one mapping between the coordinates in one space and those in another such that points in the two spaces that correspond to the same anatomical point are mapped to each other."* This definition is valid for defining most 2D-2D or 3D-3D registrations, however, it is not valid for defining a registration between a 2D projection image and a tomographic image for the following reason. Figure 1.1 shows how, if a pin-hole projection model is used, the mapping from 2D to 3D is one-to-many and the mapping from 3D to 2D is many-to-one. Therefore, in a 2D-3D registration scheme a one-to-one mapping does not exist. Instead the following definition has been used in this thesis: *2D-3D registration is the determination of a projection mapping, from a 3D to a 2D coordinate system such that points in each space which correspond to the same anatomical point are mapped*

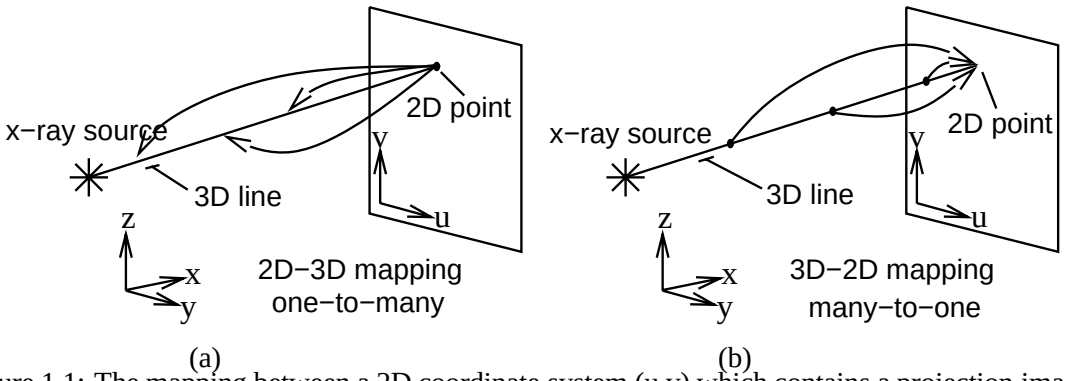


Figure 1.1: The mapping between a 2D coordinate system (u,v) which contains a projection image and a 3D coordinate system (x,y,z) can either be (a) one-to-many if the mapping is from a 2D point to a 3D line, or (b) many-to-one if the mapping is from a position along a 3D line to a 2D point.

to each other. In order to keep the definition succinct the corresponding transformation, from a 2D point to a 3D line, has not been included. However, this transformation is implicit in the above definition.

It should be noted that this definition of 2D-3D image registration is specific to the situation where the 2D image is a projection image. For 2D-3D registration where the 2D image is an ultrasound slice or a single CT slice and the 3D image is a volumetric image then the definition given by Maurer *et al.* [88] should be used.

1.2 Aim and clinical motivation

The aim of the work described in this thesis is to develop an algorithm to accurately register a pre-operative CT image to an intra-operative fluoroscopy image to aid needle or stent placement during an interventional procedure. The clinical motivation behind this aim is as follows. There are a number of interventional procedures where the initial clinical condition was diagnosed and the procedure was planned with the aid of a CT scan, however, the procedure is carried out using fluoroscopic guidance as it allows real-time imaging. During the procedure the CT slices can be viewed using a light box. However, in order to use information from the CT image during the procedure the interventionist must mentally relate information in the CT scan with information from the fluoroscopy image. By carrying out a 2D-3D registration between the pre-operative CT volume and the intra-operative fluoroscopy image it should be possible to transfer information between the 2D and 3D images much more accurately than is currently achieved. This should enable the interventionist to make more use of the pre-operative image during the procedure,

which should in turn increase the speed, accuracy and safety of the procedure.

There are also a number of procedures which are currently carried out using a combination of CT and fluoroscopy guidance which could, with the aid of a 2D-3D registration algorithm, carry out a larger percentage of the procedure under fluoroscopy guidance. Also, there are some procedures which are currently carried out using CT guidance which, with the aid of a 2D-3D registration algorithm part of these procedures could be carried out using fluoroscopy guidance. In each case, the move away from CT guidance and towards fluoroscopy guidance enables the "real-time" nature of fluoroscopy to be used. This should increase the speed and accuracy of needle or stent placement, while reducing costs.

1.3 Summary of thesis

This thesis is divided into four main sections:

1. Clinical and technical background
2. A Description of the registration algorithm
3. Validation of the algorithm
4. Accounting for spinal deformation

Each of these sections are summarised below.

1.3.1 Clinical and technical background

Chapter 2 describes the clinical motivation behind the work described in this thesis. Previous clinical applications of 2D-3D image registration algorithms are outlined. Three interventional procedures, which I propose could benefit from the use of a 2D-3D registration algorithm, are then described. These are endovascular treatment of abdominal aortic aneurysms, percutaneous laser nucleotomy and percutaneous ablation of liver metastases. In each case the current clinical protocol is outlined and then a description is given of how the 2D-3D image registration algorithm could be used to improve the speed and accuracy of the procedure.

The technical background behind the work described in this thesis is discussed in chapter 3. This chapter initially describes a framework for 2D-3D image registration, breaking down the registration process into four main stages; information extraction, similarity measures, search space

and search strategy [17]. The remainder of the chapter provides a critical analysis of eight 2D-3D image registration algorithms. These algorithms are split up into two groups; feature based algorithms and intensity based algorithms. Feature based algorithms calculate the optimum transformation between the two images by comparing specific features which have been extracted from each image. Intensity based algorithms calculate the optimum transformation between the two images by comparing voxel or pixel intensity values in each image. Each registration algorithm is described and their accuracy, robustness, speed and how applicable they are to a clinical environment is discussed. The conclusions to chapter 3 explain why I have chosen to develop an intensity based registration algorithm.

1.3.2 Description of the registration algorithm

The description of the 2D-3D registration algorithm is split up into two chapters. Chapter 4 describes the information extraction step which extracts from the CT and fluoroscopy images the information on which the registration is based. Chapter 5 discusses the other components which combine to make up the registration algorithm, namely similarity measures, search space and search strategy.

I have chosen to produce an intensity based registration algorithm. Therefore, the information extraction step involves processing the voxels in the CT volume and the pixels in the fluoroscopy image so that they can be compared by an intensity based similarity measure. Most of the work carried out in the information extraction step is the production of digitally reconstructed radiographs (DRRs) from the CT volume. DRRs were produced by projecting rays through the CT volume (as shown in fig. 1.2) and integrating the Hounsfield numbers along the ray lines. The aim is to produce a DRR which looks as similar as possible to the fluoroscopy image. To improve the similarity between the DRR and fluoroscopy image it has also been found advantageous to apply some image processing steps to the fluoroscopy image.

Chapter 5 describes seven intensity based similarity measures, most of which have been previously used for image registration. Each measure is described and then the ability of each measure to achieve an accurate 2D-3D registration between an interventional fluoroscopy image and a CT image is discussed. The chapter then describes the search space over which the algorithm is expected to perform. Finally the chapter discusses the search strategy used to optimise a given similarity measure. The discussion is split up into two sections, firstly how the rigid body parameters are altered to achieve a registration and secondly the choice of axes about which rotations and translations occur.

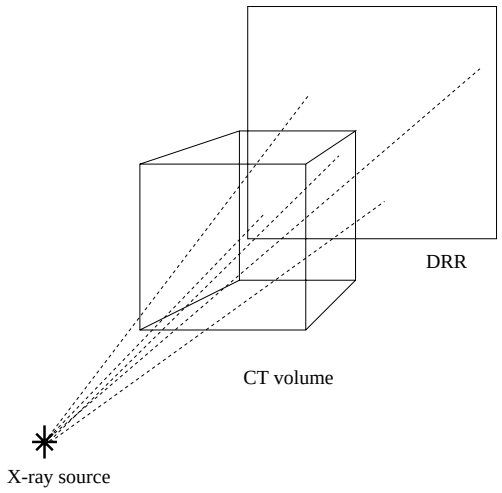


Figure 1.2: The formation of a DRR by casting rays through the CT volume

1.3.3 Validation of algorithm

A validation of the 2D-3D algorithm is carried out in Chapters 6 and 7. Chapter 6 validates the registration algorithm using images of a lumbar spine phantom. Final registration positions are compared to a "gold-standard" registration which has been accurately calculated using fiducial markers. The seven similarity measures, described in chapter 5, were each used in the registration algorithm and the accuracy and robustness of the different similarity measures was investigated. The effect of a number of factors on the performance of the similarity measures was investigated. Experiments were carried out to determine the effect of overlying soft tissue structures and interventional instruments on the phantom fluoroscopy image. Also, the effects of altering CT slice thickness, of altering the out-of-plane translation direction and the effect of including non-physical rays in the DRR were investigated.

Chapter 7 continues the validation of the algorithm but uses clinical images from aortic stenting procedures. It was not possible to calculate an accurate "gold-standard" registration for these images and so *k-fold* cross validation [66] was used to estimate the registration accuracy.

In *k-fold* cross validation a set of data is split up into *k* subsets of approximately equal size. All except one of these subsets are used as a set of training data and the other subset is used as a set of test data. The training data is used to calculate a particular set of parameters which are then used to test each element of the test data set. This process is repeated *k* times where a different subset of data is left out each time. My method was to register to a number (n) of different vertebrae. The registration results for one vertebra were removed to form a test data set and then the registration results to the other vertebrae were used to form a training data set. This process was repeated n

times removing the results from a different vertebra each time.

Care had to be taken in the design of the validation experiment, so that any relative movement of the vertebrae between the pre-operative and intra-operative images did not effect the results. The method employed used two fluoroscopy images. Registrations were carried out individually to each vertebra in each fluoroscopy image. The registrations to the same vertebra in both fluoroscopy images were used to calculate the transformation of the fluoroscopy set between when the two fluoroscopy images were acquired. Assuming that no movement of the vertebrae occurred between the acquisition of the two fluoroscopy images, then the transformation of the fluoroscopy set should be invariant to any relative movement of the vertebrae between the pre-operative and intra-operative images. This is a valid assumption as the patient is not moved between the acquisition of the two fluoroscopy images.

The final part of chapter 7 describes an experiment to register a pre-operative CT volume to an intra-operative fluoroscopy image from a percutaneous laser nucleotomy procedure. A registration was also carried out between the pre-operative CT volume and a pre-operative magnetic resonance (MR) image using a 3D-3D registration algorithm based on normalised mutual information [112, 116]. By combining these two registrations it was possible to overlay information from the pre-operative MR image onto the intra-operative fluoroscopy image.

1.3.4 Accounting for spinal deformation

In chapter 8 the 2D-3D registration algorithm is combined with a deformation algorithm. The objective is to deform a pre-operative CT image so that it more accurately represents the intra-operative scene as viewed in a fluoroscopy image. The chapter describes the deformation algorithm designed by Little *et al.* [75]. This algorithm allows rigid bodies to be incorporated into a non-linear deformation based on a radial basis function. The chapter then describes an improvement to this algorithm which results in more realistic deformations close to the rigid bodies. The deformation algorithm is used to warp a pre-operative CT volume based on information obtained using the 2D-3D registration algorithm. Images from an aortic stenting procedure were used. Four registrations were carried out, each to a different lumbar vertebra. The registration results were used to calculate the amount of relative vertebral movement which has occurred between the pre-operative and the intra-operative images. This information is then input into the deformation algorithm where it is used to define the rigid body motion of each vertebra. The four vertebrae are transformed as individual rigid bodies and the rest of the image is allowed to deform based on the movement of the vertebrae and a number of point landmarks.

Chapter 2

Clinical Background

This chapter describes the clinical motivation behind the work in this thesis. It attempts to answer the following question: why design a 2D-3D image registration algorithm for use in image guided interventions? The chapter begins by describing three different fields; multi-modality image registration, image guided surgery and interventional radiology. The work described in this thesis aims to combine the ideas from these three fields to help solve some clinical problems. The current clinical uses of 2D-3D image registration algorithms are discussed. Then three interventional procedure are described and I discuss how these procedures could be aided by using a 2D-3D image registration algorithm.

2.1 Multi-modality image registration, image guided surgery and interventional radiology

This section initially outlines multi-modal image registration, image guided surgery and interventional radiology. It then discusses how it should be possible to use ideas from image guided surgery and from multi-modal image registration to help solve some clinical problems in the field of interventional radiology.

2.1.1 Multi-modality image registration

There is an increasingly wide range of imaging modalities available to clinicians. These different modalities are often complementary, displaying different features within the patient. For example, CT shows excellent bony detail, MR is often the preferred modality for visualising soft tissue structures and positron emission tomography (PET) images display functional rather than anatomical information.

Often a single modality is not sufficient for patient diagnosis, so a second modality is required. In the past these were printed on a film and viewed with a lightbox. If the clinician wanted to compare the spatial positions of information from two modalities then they would use a "poorly described, yet generally understood, visual alignment system" [88]. In recent years the use of dedicated medical image computer workstations to view medical images has become more widespread and a range of image registration algorithms have been developed to calculate the point to point correspondence between different images of the same patient. These registration algorithms fall into two main categories; the first optimises a similarity measure based on corresponding features in the two images; the second optimises a similarity measure based on the overlapping voxel intensities from each image (good reviews on these registration techniques can be found in [76, 88, 123]). The first image registration algorithms were feature based and required the identification of corresponding landmarks in both of the images. Feature based approaches tend to require more user interaction than the voxel based methods. The recent trend in voxel based methods has been to use a similarity measure based on information theory called mutual information [27, 124]. Algorithms using a mutual information similarity measure [27, 113] were found to be amongst the most accurate in a retrospective registration trial undertaken by Vanderbilt University [137]. This investigation compared the accuracy of several 3D-3D multimodal registration algorithms to register CT and PET images of the brain to corresponding MR images.

2.1.2 Image guided surgery

Registration techniques have also been used to relate an image coordinate system to physical world coordinate system to help guide surgical procedures. These coordinate systems are shown in figure 2.1. The patient is registered to the pre-operative image, i.e. the transformation $\mathbf{T}$ (see figure 2.1) is calculated, and then the spatial position of features in the pre-operative modality can be accurately related to world positions in the operating theatre to help guide the surgical procedure.

To register physical space to a pre-operative volume it is necessary to define a physical world coordinate system within which 3D positions can be accurately determined. One method of achieving this is to use a stereotactic frame. A stereotactic frame is a mechanical device which is used to accurately position surgical instruments in 3D space. The frame is rigidly attached to the patient's skull and is used to define a world coordinate system [88]. Other methods to define a world coordinate system are to use mechanical [105], electromagnetic [10], ultrasonic [2], optical [20] or video-based [26] tracking.

To register image space to physical space typically requires corresponding features to be

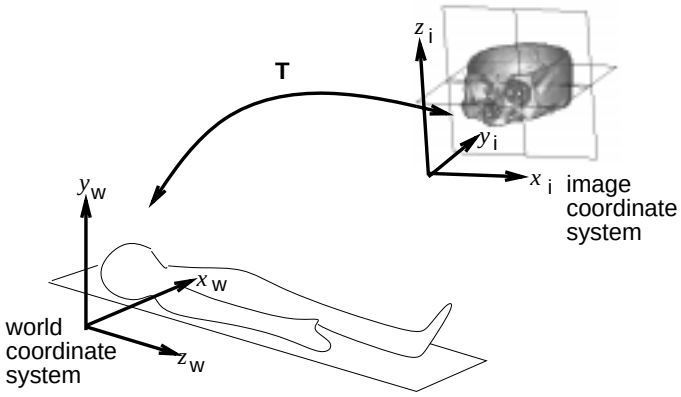


Figure 2.1: The coordinate systems in a typical image guided surgery system. The aim is to find the transformation $\mathbf{T}$ which relates the world coordinate system (x_w, y_w, z_w) to the pre-operative image coordinate system (x_i, y_i, z_i). Once $\mathbf{T}$ is known the surgeon can accurately relate positions in the surgical scene to positions in the pre-operative image.

identified in each coordinate system. If a stereotactic frame is used the pre-operative images of the patient are acquired with the frame attached. Features on the frame, usually six vertical and three diagonal rods arranged in three N shapes, are automatically found in the images. The position of these rods in image coordinates are registered to the known rod positions in the world (frame) coordinate system. This enables the position and orientation of each pre-operative image slice to be determined with respect to the world coordinate system [88]. If an optical localisation device is used a commonly employed system is to attach fiducial markers to the patient before the pre-operative images are acquired. The markers are designed so that they can be accurately localised in the pre-operative image and also in the world coordinate system. The two sets of points are then used to calculate a rigid body transformation between the image and the world coordinate systems. Highly accurate registration has only been validated using bone-implanted markers [89]. Less invasive techniques, such as skin-attached markers or anatomical positions,have been found to be less accurate [127]. Work continues to develop less intrusive methods such as the use of a bite block [39, 120] on which to attach a stereotactic frame or fiducial markers, or the use of intra-operative video images [25]. These methods have yet to be shown to be as accurate as bone implanted markers.

Once registration has been achieved, information from the pre-operative image can be used to help guide the surgical procedure. The most common method when using an optical localiser is to display the 3D position of the localiser within the pre-operative image, usually by reformatting the

pre-operative volume to display three orthogonal planes, where the position of the tip of the pointer is at the intersection of these three planes. The planes can also be oriented so that two planes are parallel and one is normal to the direction of the pointer. It is also possible to transfer information in the other direction and display features from the pre-operative modality into the surgical scene. Such a system is being developed by Edwards *et al.* [39], where segmented features from a pre-operative image are projected into the eye-pieces of a surgical microscope to augment the surgeons view.

2.1.3 Interventional radiology

Interventional radiology is a relatively new clinical discipline which has developed over the last three and a half decades. The use of radiographic images to guide a needle to a specific anatomical area goes back a number of years, but it was only with the development of fluoroscopic image intensifiers in the 1950's that interventional radiology began to establish itself. Prior to the 1950's direct vision fluoroscopy was used. This involved a radiologist observing a phosphor screen (through a sheet of lead glass) which converts the pattern of x-rays, which have passed through the patient, into visible light. In order to keep the radiation dose to the patient at an acceptable level the phosphor screen was viewed in a darkened room. However, even though the spatial resolution and contrast of the fluoroscopy image was nearly as good as conventional radiography, the brightness of the screen was so low that the eye was only able to perceive approximately 1/10th of the resolution and contrast on the screen [41]. A fluoroscopy image intensifier uses electrical energy to increase the brightness of the fluoroscopy image so that it can be viewed in a normally lit room or by a video camera. This removed the requirement for dark adaptation to dim fluoroscopic screens and allowed the radiologist to observe the full resolution and contrast available in the fluoroscopy images. More details on the production of fluoroscopy images are given in section 4.1.1.

Other imaging modalities, such as ultrasound, CT and MR, have also been used to guide interventional procedures. The use of cross sectional imaging extended the scope of interventional radiology to solid organs, whereas with fluoroscopic guidance only hollow organs could be tackled, as they could be visualised using contrast media. Fluoroscopy and ultrasound are by far the most common interventional imaging modalities because they allow the position of interventional instruments to be visualised in real-time. The recent introduction of interventional MR machines and CT fluoroscopy has made these modalities more suitable for interventional procedures by allowing faster imaging and allowing better access to patients. However, due to the high cost

of these machines the vast majority of interventional procedures are currently carried out under fluoroscopic or ultrasound guidance.

Interventional radiology is described by Adam as “a discipline which uses imaging guidance to effect treatment” [1]. A number of interventional procedures involve intravascular catherization. These include the infusion of chemical agents into tumours, placement of radioactive substances, embolization of vessels, angioplasty to widen narrowed vessels and the placement of filters [99]. There are also a large number of procedures which involve the insertion of needles and/or catheters into the thorax, biliary tree, urinary system, retroperitoneal tissue and skeleton. These techniques are performed by interventional radiologists rather than surgeons as one of the major factors involved is the ability to interpret the images used to guide the procedures. Many of these techniques offer a therapeutic alternative to open surgery, which can be performed in a shorter time, with a lower morbidity and at a reduced financial burden than their surgical counterparts [99]. Also, because interventional procedures are generally less invasive than the surgical alternative, these procedures can be carried out on patients who previously, due to associated medical problems, would not have been considered as suitable candidates for surgery.

2.1.4 Combining these ideas

The aim of the work described in this thesis is to combine ideas from image guided surgery with multi-modal image registration techniques to allow information from a pre-operative modality to be used during interventional procedures. The long term aim is to produce a system similar to image guided surgery but for the field of interventional radiology.

In conventional open surgery the surgeon can directly observe the surgical scene and this visual information, along with tactile feedback, is used to guide the operation. In interventional radiology the interventionist is unable to directly view the anatomy on which the procedure is being carried out. Instead an imaging modality is used which, along with tactile feedback, provides the real time information used to guide the procedure. Image guided surgery attempts to relate the spatial position of information from a pre-operative modality to a physical world coordinate system which encompasses the surgical scene. The aim of the work described in this thesis is to use image registration techniques to relate the spatial position of information from a pre-operative modality to the spatial position of information in an intra-operative modality.

In one respect the registration problem posed in this thesis is easier than the registrations carried out in image guided surgery, because a localisation device is not required to map out a world coordinate system. The registrations in this thesis will be between two digital images and

so the two coordinate systems are already defined. However, the image registration problem is different to a 3D-3D multi-modal image registration as the registration is between a 3D image and a 2D projection image. This change in dimensionality can create additional problems. Three dimensional images contain much more spatial information than 2D projection images. There is, therefore, less spatial information common to both images in 2D-3D registration than in 3D-3D registration. In particular the 2D images contain much less spatial information on the position of features perpendicular to the imaging plane, compared to the position of features parallel to the imaging plane. Consequently the registration accuracy perpendicular to the imaging plane is expected to be less than in the other degrees of freedom.

The proposed clinical applications for the 2D-3D registration algorithm are also in very different parts of the body compared to image guided surgery procedures, which tend to concentrate on neurological or skull base surgery. In image guided neurosurgery the skull and brain are often assumed to behave as a rigid body, though recent studies have shown that shifts can occur under a large crainotomy [58, 86]. The clinical applications proposed in section 2.3 are all in the abdominal region where soft tissue deformation will occur and vertebrae will move relative to each other.

2.2 Current clinical uses of 2D-3D image registration

Registration of a tomographic image and an x-ray projection image has been proposed to help in a number of clinical areas. Image registration between CT and simulator or portal images has been put forward as an aid to patient placement for radiotherapy planning and treatment verification [5, 9, 45, 94]. Other papers have proposed using 2D-3D image registration to help in neurointerventions [42, 65, 77, 78], spinal surgery [118, 131], hip replacement [51] and aortic stenting procedures [131]. Another application put forward is to register the position of post-operative radiographs to a pre-operative CT volume to verify the position of intra-cerebral electrodes [73].

At present 2D-3D registration algorithms are still under development and I know of none which are in widespread clinical use. Some of the above algorithms have been demonstrated under clinical conditions for applications in neurosurgery [25], hip replacement [51] and aortic stenting [16]. The area in which 2D-3D registration appears closest to playing an important clinical role is in the field of radiotherapy. Radiotherapy could benefit greatly from 2D-3D registration by aligning the CT treatment plan either with the simulator image or with a portal image. Simulator images are standard radiographic images taken prior to treatment using x-ray apparatus which is designed

to have the same geometry as the treatment machine. Portal images are taken during radiotherapy treatment by placing an image device (either a portal film or an electronic portal imaging device) to record the amount of transmitted radiation at the beam exit side of a patient. Registration to a simulator image could increase the accuracy of patient positioning, while registration to a portal image could be used to verify patient position during therapy. At present, the main method for treatment verification in radiotherapy is a subjective process based on the visual comparison of portal and simulator images. A recent study has shown that there can be large inconsistencies between different clinicians assessments of the same pair of portal and simulator images [11]. Because of these inconsistencies there is a large amount of interest in developing more objective methods to compare simulator and portal images. These range from providing computer tools in electronic view boxes [14], to automated 2D-2D registration algorithms [37, 46, 61, 101, 129]. These techniques aim to increase the speed and improve the accuracy of portal image verification. However, these techniques are two-dimensional and so cannot calculate the out-of-plane errors. A 2D-3D registration enables patient positioning errors in all six degrees of freedom to be calculated. Lujan *et al.* [80] have extended a 2D-2D registration scheme to also calculate out-of-plane errors and a number of papers report full 2D-3D registration algorithms [5, 9, 45, 94].

2.3 Proposed interventional procedures for 2D-3D image registration

The majority of proposed clinical applications for 2D-3D registration algorithms have been to aid radiotherapy or surgical procedures. The clinical motivation behind the work described in this thesis was to aid image guided interventions. In particular the aim is to improve stent or needle placement in three interventional procedures; endovascular treatment of abdominal aortic aneurysms, percutaneous laser nucleotomy and percutaneous ablation of liver metastases. Each of these procedures are described and I then discuss my intended method for improving the speed and accuracy of stent or needle placement by using a 2D-3D image registration algorithm. The aim in each case is to combine the excellent 3D spatial information and soft tissue contrast from CT images with the high temporal resolution and dynamic nature of fluoroscopic imaging. The alternative to these interventions (where there is one) is a surgical procedure which typically has a number of disadvantages, namely, the need for the use of a general anaesthetic, a longer hospital stay and more damage to the surrounding tissue.

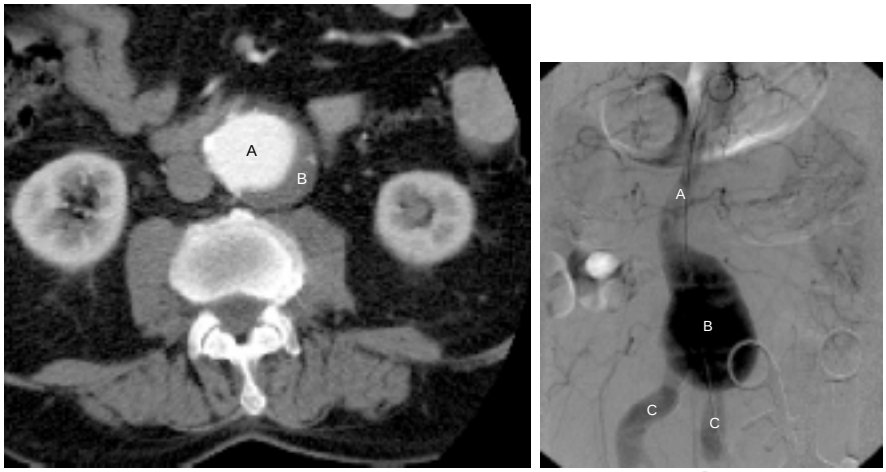
2.3.1 Endovascular treatment of abdominal aortic aneurysms

Recent advances have introduced methods to treat abdominal aortic aneurysms using an endovascular approach. This approach has a number of potential advantages over conventional surgery. These include a reduction in morbidity and mortality rates, due to a reduction in the use of anaesthesia, surgical dissection, blood transfusion and aortic clamping. Other advantages are shorter hospital stays and reduced costs [93]. The endovascular approach may also allow for the treatment of patients who were not candidates for standard resection because of associated medical conditions [121]. At present the reported operative mortality compares favourably with the mortality rates in standard open surgical repair [121]. Also, a recent clinical trial showed that the length of stay was approximately half that of a standard surgical repair and although at present the costs were similar, these may be expected to change in the future when procedures are simplified and conversions to open surgery, due to complications during the procedure, are kept to a minimum [82].

The present procedure is as follows: a pre-intervention CT angiogram and digital subtraction angiography (DSA) image are acquired. These are used to assess the patient's suitability for the procedure and to select the correct stent type and size. The CT volume is particularly important in assessing the size of the aneurysm as it can visualise both the patent lumen and thrombus, see figure 2.2(a). In the past, when pre-operative imaging consisted of only a DSA image, see figure 2.2(b), the size of the aneurysm tended to be underestimated due to thrombus adherent to the walls of the aneurysm [109]. The stents used are of two types, either a tube or a bifurcated configuration, the latter configuration is used if the distal end of the aneurysm is very close to the aortic bifurcation.

The procedure is carried out under general or epidural anaesthesia. A small incision is made in one of the femoral arteries and a guide wire is advanced up the aorta to the level of the diaphragm. If a bifurcated stent is to be deployed both femoral arteries are used. Fluoroscopic guidance is used to pass a delivery sheath (containing the stent) up the guidewire and into the aorta. The positioning of the stent is determined by referring to an intra-operative angiogram. When in the correct position the stent is deployed from the delivery sheath, see figure 2.3. After deployment a DSA image is acquired to confirm good flow through the stent and the exclusion of the aneurysm sac. The objective is to protect the aneurysm from the outward force created by the arterial blood pressure.

Accurate placement of the stent is very important. The renal arteries must not be covered, as this will cut off the supply of blood to the patient's kidneys. If a tube configuration is used, then



(a) (b)

Figure 2.2: Images from aortic stenting procedure, (a) axial slice through pre-operative CT angiogram, showing (A) patent lumen and (B) thrombus in aneurysm, (b) intra-operative digital subtraction angiography image showing patent lumen in aorta (A), aneurysm (B) and iliac arteries (C).

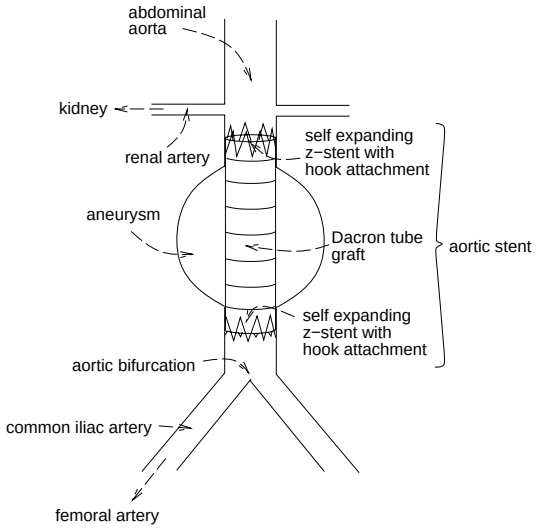


Figure 2.3: A correctly deployed aortic stent (tube configuration) and the surrounding vascular anatomy. The stent is deployed below the renal arteries, above the bifurcation of the aorta and the whole of the aneurysm is covered.

the stent must not be deployed below the bifurcation of the aorta as this may obstruct the blood flow into one of the iliac arteries, and the whole aneurysm must be covered so as to exclude the aneurysm sac from the blood flow. Therefore, in order to place the stent correctly the interventionist needs to be able to visualise the proximal and distal necks of the aneurysm, the renal arteries and the bifurcation of the aorta. Plain fluoroscopy cannot visualise these features. The patent lumen of the aneurysm and the arteries can be visualised using an intra-operative DSA image, see figure 2.2(b). However, it is important to know if the aneurysm neck, as seen in the DSA image, is formed by the aortic wall and is not thrombus within the aneurysm sac [90]. CT reliably images the patent lumen and thrombus [109] and, therefore, it could be advantageous to use information from the CT image in positioning the stent. By registering the pre-operative CT volume to the intra-operative fluoroscopy image, information on the patent lumen and thrombus could be projected from the CT volume and overlaid onto the fluoroscopy image. Information could also be transferred in the other direction. By carrying out two 2D-3D registrations, to two fluoroscopy images taken from different views, corresponding 2D points in each fluoroscopy image (such as the tip of the stent catheter) could be displayed as a 3D point in the pre-operative CT volume. These methods could be used to ascertain whether the stent was positioned correctly with respect to the patent lumen and thrombus. In discussion with the interventionists at Guy's Hospital we estimate that the required registration accuracy is approximately 1mm. This figure for the required registration accuracy (and the accuracy figures quoted for the other procedures in the subsequent sections) is a rough estimate based upon the interventionists initial ideas of how a 2D-3D registration system could work in clinical practice. I expect that these required accuracy figures will vary greatly between individual clinical cases. More information on the required accuracy could be obtained by investigating individual clinical cases, however, this goes beyond the current scope of research discussed in this thesis. Soft tissue deformation is not expected to introduce registration errors because the registration will be based on information from the vertebrae and the aorta is fairly rigidly attached to the spine.

2.3.2 Percutaneous laser nucleotomy

Percutaneous laser nucleotomy (PLN) is a relatively new interventional procedure for the treatment of lower back pain caused by a prolapse of an intervertebral disc applying pressure to nerve roots. The procedure involves inserting a needle into the centre of the intervertebral disc. A laser is used to vaporise part of the centre of the disc. This reduces the pressure within the disc, which in turn reduces the pressure on the surrounding nerve roots. The first PLN procedure was carried

(a) (b)

Figure 2.4: Images from a percutaneous laser nucleotomy procedure, (a) interventional fluoroscopy image showing 2D position of needle, (b) axial slice from an interventional CT volume showing (A) 3D position of needle and (B) root of first sacral nerve.

out by Ascher and Choy in 1986 [23]. The procedure has a number of potential advantages over conventional surgical approaches: it can be carried out on an outpatient basis at a reduced cost and has a shorter treatment time [23, 74]. It also offers another option to patients who have resisted open surgery [74].

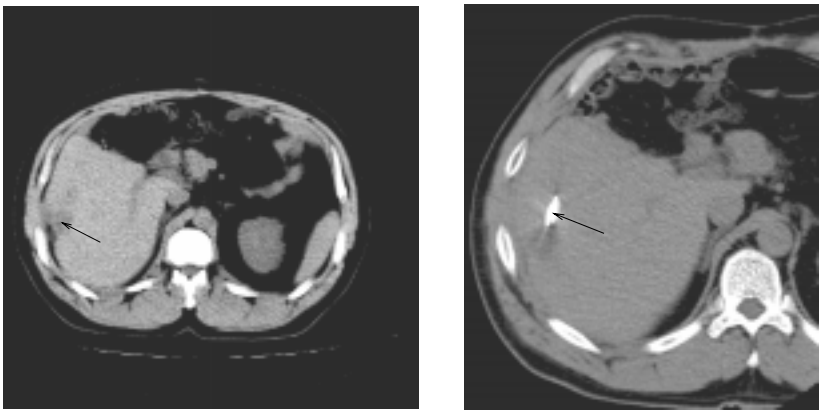
The initial procedures were carried out using biplane fluoroscopic imaging [23]. Interventional CT, MR and microendoscopy have also been used to guide the procedure [50]. At Guy's Hospital the procedure is carried out using a combination of CT and fluoroscopic guidance [32]. The present procedure is as follows: a pre-intervention MR and/or CT has been acquired for diagnostic purposes. The patient is placed on the CT table in the prone position under local anaesthesia and light sedation is established. Heavier sedation is not used as the patient must be awake enough to respond if the nerve root is irritated by the advancing needle or probe. Once in position a preliminary CT scan is acquired to determine the optimum point of entry and the desired needle track. Fluoroscopy is then used to guide the needle into the intervertebral disc in real time so that the needle tip penetrates the annulus fibrosis, see figure 2.4(a). Intermittent CT scans are also taken to localise the tip of the needle in 3D, see figure 2.4(b). Once in position an optical fibre is threaded down the needle and laser energy is delivered to the disc to vaporize part of the nucleus pulposus. The resultant gas is allowed to escape back up the needle. The small change in mass of the nucleus pulposus, because it is within a confined area, results in a large decrease in pressure within the disc, which should in turn reduce the pressure on the nerve roots leading to an improvement in symptoms. Further CT images are taken to monitor the gas production while the laser energy is delivered. When the procedure is finished, the needle and fibre are withdrawn and a dressing

applied. Success rates of 78%, 72% and 60% have been reported by Choy *et al.* [23], Liebler [74] and Cowling *et al.* [32] respectively.

According to Liebler [74] the two most important aspects of this procedure are patient selection and needle placement. For guiding PLN procedures CT images have a number of advantages over fluoroscopy images. If there are vital structures, such as nerves, in high risk areas then CT is the "gold-standard" of guidance technology [50]. CT allows the visualisation of structures such as nerve roots (see figure 2.4(b)) or the thecal sac which cannot be observed on fluoroscopy images. Therefore, the risk of injury to these structures is reduced by employing CT guidance. On the other hand, fluoroscopy is a more dynamic imaging modality and so allows faster positioning of the needle. The current clinical protocol aims to combine these two imaging modalities. The position of the nerve roots are identified using the initial CT scan and this information is then mentally related to the fluoroscopy image by the interventionist using their knowledge of anatomy. The use of a 2D-3D image registration should make the transfer of information between the CT and fluoroscopy images more accurate. An accurate 2D-3D registration would make it possible to overlay structures of interest from the CT scan, onto the fluoroscopy image. Also, if two fluoroscopy views are taken of the interventional scene, then it should be possible to show the position of the needle tip within the initial or pre-operative CT volume. If other pre-operative modalities are available (such as MR) and they can be registered to the pre-operative CT scan using a 3D-3D registration algorithm [112], then they can also be used to aid the needle placement. I propose that by using a 2D-3D registration algorithm the accuracy, speed and safety of needle positioning will be increased. In discussion with the interventionists at Guy's Hospital the required registration accuracy for PLN has been estimated to be 1mm. Vertebral movement may be an issue in the registration. However, the amount of deformation between a vertebra, an adjacent disc and the lumbar nerves is expected to be small.

2.3.3 Percutaneous radio-frequency ablation of liver metastases

The spread of metastatic disease to the liver is one of the most common causes of death in patients with colon cancer (13,000 deaths per year [3]). Recent studies have shown that a significant increase in survival time and even a cure is possible if all the liver metastases are removed [110]. Surgical resection has been the only established treatment offering the potential for cure, however, only a minority of patients are surgical candidates, 5% – 20% of patients whose primary tumour is colorectal carcinoma [34]. Therefore, there are a large number of patients for whom there is no effective treatment [3]. Recently a number of minimally invasive techniques for ablation



(a) (b)

Figure 2.5: Axial CT slices from percutaneous liver metastases procedure, (a) initial CT scan where arrow shows the position of the target metastasis, (b) CT slice showing the position of the tips of the RF probe (marked by the arrow) Note, because an oblique trajectory was used only a short section of the probe is seen in each axial slice.

of liver metastases have been proposed including percutaneous introduction of alcohol, radio-frequency heating, microwave heating and laser photo-coagulation [3, 34, 110, 125]. Percutaneous ethanol injection is the treatment of choice for unresectable hepatocellular carcinoma. Metastases are more resistant to the uniform distribution of injectable agents than hepatocellular carcinoma due to their firm texture and therefore, radio-frequency (RF), laser or microwave therapy may be more appropriate [34]. Recent studies have shown that both RF and laser ablation appear to be simple, safe and potentially effective methods for the treatment of colorectal liver metastases [3, 110, 125]. The treatment can be carried out on an outpatient basis at a significantly lower cost than surgical resection and can be offered to a number of patients for whom no treatment was previously available.

Planning requires accurate delineation of all metastatic disease in the liver. Treatment is most effective when there are less than 3 metastatic growths, each less than 40mm in diameter [128]. In our institution therapy is currently delivered under CT guidance. The interventionist decides on a trajectory for a guide needle after examination of the contrast enhanced CT (see figure 2.5(a)). The needle is inserted percutaneously and individual CT slices are acquired with the CT gantry oriented so that the needle lies within the slice plane. The needle is gradually advanced towards the metastasis and CT slices are retaken as necessary to assess progress towards the target. The RF probe is a multi-probe array consisting of three needles, each of which have an RF electrode at their tip, see figure 2.6. The three needles are inserted into the metastasis along approximately the same path as the guide needle. The RF probe causes tumour necrosis by heating the tumour cells to approximately 50^{o}C. The heating is carried out for approximately twelve minutes and then the probe and needle are withdrawn and a dressing applied.

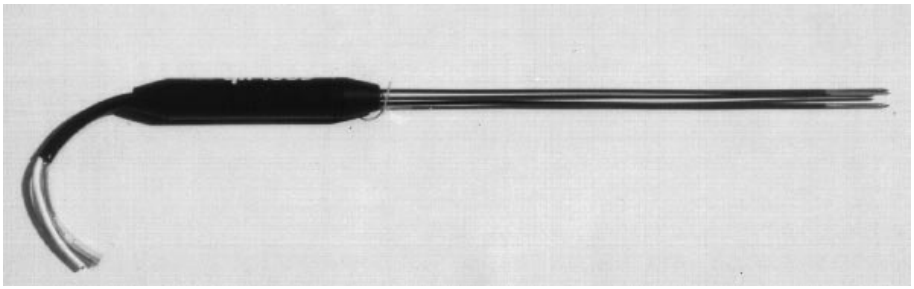


Figure 2.6: Radionics radio frequency probe. A cluster of three needles is used to increase the size of the treated volume.

Accurate placement of the needles in this procedure is important to ensure that all of the tumour is destroyed. A number of imaging modalities have been used to guide percutaneous ablation of liver metastases procedures, these include ultrasound [3, 110], CT [125] and MR [126]. As described above, the present procedure at Guy's Hospital is carried out using CT guidance. The procedures can take a couple of hours to perform and the majority of this time is taken up in accurately placing the needle within the metastasis. Speeding up the procedure could allow for the treatment of more metastases per session and reduce patient discomfort. I am proposing that this procedure could be carried out using 2D-3D image registration and a combination of CT and fluoroscopic guidance. An initial CT scan would be taken from which the position and trajectory of the guide needle would be determined. The patient would then be moved out from the CT gantry and the needle placement would continue under fluoroscopic guidance. The 2D-3D registration algorithm would be used to register the fluoroscopy image to the initial or pre-operative CT scan and then the position of the target metastasis could be projected onto the interventional fluoroscopy image. Or, if two fluoroscopy images were taken from different views then the position of the tip of the needle could be localised within the initial or pre-operative CT volume. If other pre-operative modalities, such as PET images were available, and they were registered to the CT volume using a 3D-3D registration algorithm [112], then these images could also be used to aid the needle placement.

Fluoroscopy guidance allows a wide variety of views to be used, while the slice orientation of a CT scanner is limited by the amount of gantry tilt available ($\pm$ 30° on the Philips TOMOSCAN SR7000). An oblique needle trajectory is often required for the treatment of liver metastases and sometimes the CT gantry cannot be tilted by a large enough angle to obtain slices in the same plane as the needle (see figure 2.5(b)). This results in only a section of the needle appearing in each CT slice making it difficult to localise the needle tip and so increasing the time taken to complete the procedure.

Soft tissue deformation may cause a problem if a pre-operative modality is used to help guide a percutaneous liver metastases procedure. The liver is a deformable organ, as seen by the shape it adopts on the butcher's slab. Also, the liver is directly below the diaphragm and so will deform due to the breathing motion of the patient. The pre-operative modality will, therefore, only be valid for a single instant during the breathing cycle of the patient. A preliminary study was carried out by Barden [6] to ascertain the expected size of liver deformation. This study showed that although the range of liver motion during normal shallow breathing is quite large (12mm), the volunteers were able to hold their breath (and hence also position their liver) to a fairly high precision; a standard deviation of 3.0mm. Because the size of a metastasis is typically a few cm across, fairly large registration errors can be tolerated. The registration accuracy required for the treatment of liver metastases is expected to be approximately 5mm. Therefore, if patients can reproduce their breathing as accurately as the volunteers in Barden's study [6] then deformation may not be a problem. Another solution to soft tissue deformation is to deform the pre-operative volume so that it more accurately represents the intra-operative scene. This could use information from the position of features which will move with the liver and are visible on the fluoroscopy screen such as the ribs or lung boundary.

2.4 Summary

The clinical motivation for the work in this thesis has come from three main areas. The intention is to combine ideas from the field of image guided surgery with multi-modal image registration techniques to aid image guidance during interventional procedures. By registering pre-operative CT volumes to intra-operative fluoroscopy images the aim is to combine the excellent three dimensional spatial information and soft tissue visualisation from CT with the real-time imaging capabilities of fluoroscopy. The work in this thesis is directed towards aiding three interventional procedures; endovascular treatment of abdominal aortic aneurysms, percutaneous laser nucleotomy and percutaneous ablation of liver metastases. The hypothesis is that 2D-3D registration can aid the image guidance of these three procedures and so increase the accuracy and speed of stent or needle placement. This should help to improve the clinical outcome and reduce patient discomfort. This thesis is primarily concerned with algorithm development and validation of the 2D–3D registration algorithm. Building and validating a complete system for image guided interventions is beyond the scope of this thesis.

Chapter 3

Review of 2D-3D registration algorithms

The first 2D-3D medical image registrations were performed in the field of radiotherapy in the 1980s [21, 35]. When CT images appeared on the clinical scene their use in external beam radiotherapy soon became apparent. CT images are an ideal modality on which to plan radiotherapy treatment as they are obtained in the axial plane. However, a 2D-3D image registration was required to relate the planned beam position on a CT scan, to the actual or simulated beam position on a portal image or simulator film.

The first algorithms to register 2D and 3D images for use in surgical or interventional procedures were published in the early 1990's by Lavallée *et al.* [71]. As with 3D-3D image registration the early work in this subject concentrated on feature based registration, where the registration is based on corresponding points, lines or surfaces found in the images. The first work using an intensity based approach was published in 1994 by Lemieux *et al.* [73] and since then a number of papers have been published using intensity based approaches [45, 49].

The aim of this chapter is to review the best 2D-3D algorithms in the literature to decide what type of 2D-3D registration algorithm I should develop. The chapter begins by describing a framework for 2D-3D image registration. This framework is used in the subsequent sections to discuss eight state of the art 2D-3D algorithms in the literature. The algorithms are divided into feature based and intensity based approaches. These algorithms have been used to register images varying from computer simulated images to routinely acquired clinical images. The next section concentrates the discussion on the performance of these algorithms when registering a CT volume to one or more real x-ray projection images.

3.1 A framework for 2D-3D image registration

The process of image registration can be broken down into a number of steps. Brown [17] splits up the process of image registration into a framework of four main stages, namely, feature space, similarity measure, search space and search strategy. I have used a very similar framework to Brown [17], except I use the term information extraction, instead of feature space for the following reasons. Firstly, the term feature space may cause confusion because intensity based algorithms do not extract any features. Secondly, the term feature space may again cause some confusion because it is also used in intensity based 3D-3D image registration to describe grey-level scatter plots [28]. The following sections describe each stage of the registration process, where particular attention is drawn to the differences between 2D-3D registration and 2D-2D or 3D-3D registration.

3.1.1 Information extraction for 2D-3D registration

The information extraction step obtains the information on which the registration is based, from the raw images. A number of different sets of information can be chosen, such as pixel intensities, edges, contours, surfaces, line intersections, etc. If the information extracted is comprised of salient features then the resultant registration algorithm is said to be feature based. By salient features I refer to features which have been located at a meaningful position, such as the intersection of two lines, or a specific contour. If the information extracted is comprised of pixel intensities (or a function, such as the gradient of the pixel intensities) from the images, then the resultant registration algorithm is said to be intensity based.

In selecting the information on which to base the 2D-3D image registration algorithm care has to be taken to extract two sets of information which are compatible. In order to extract compatible sets of information the physics behind the two imaging modalities should be taken into account. This thesis is primarily concerned with the registration of fluoroscopy images with CT volumes. Regions of low pixel intensity in a fluoroscopy image are due to high attenuation materials within the imaged object lining up along the same x-ray path. Therefore, the pixel intensities within the fluoroscopy image are created by a combination of both the structure within the object and the pose of the object. Voxel intensities in the CT volume are linearly related to the attenuation coefficients within the imaged object and are not affected by the pose of the object. This is different to 3D-3D image registration where the information in both images does not depend on the pose of the object.

In 2D-3D image registration the two sets of extracted information may have different dimensionality. For example, a set of 3D points could be extracted from the CT volume and a set of

corresponding 2D points could be extracted from the fluoroscopy image. These two sets of points could then be compared using a set of projection parameters. Either the 3D points could be projected into 2D, or the 2D points could be back projected to form 3D lines. Also, in 2D-3D image registration, the extracted information may have a different dimensionality compared to the image it was created from. An example of this is the production of a 2D DRR from a 3D CT volume.

Once the information extraction step is completed, it is these two sets of information, rather than the two images, which are registered together. Therefore, care has to be taken that errors are minimised during the information extraction step.

3.1.2 Similarity measures

The similarity measure calculates a numerical value which represents how well the two sets of extracted information have been aligned. Therefore, the choice of similarity measure is greatly effected by what information is extracted. If the information extracted is comprised of salient features then, typically, a feature based similarity measure is used. Feature based similarity measures are often a physical quantity, such as the distance between corresponding points, or the minimum distance between lines. Often a large number of features are extracted and the algorithm not only has to minimise the distance between features, but must establish feature correspondence between the two images. The complexity of feature correspondence increases dramatically as the number of features increases.

If the information extracted is comprised of pixel intensities then, typically, an intensity based similarity measure is used. Examples of commonly used intensity based similarity measures are the sum of absolute differences, cross correlation and mutual information.

When registering on clinical images some of the information extracted from one image may have no counterpart in the set of information extracted from the other image. This could be due to a number of reasons, the different imaging modalities, a different field of view, or the presence of additional structures such as an interventional instrument. A good similarity measure should be able to register on the extracted information common to both images and remain invariant to any information extracted from one image which has no counterpart in the set of information extracted from the other image.

An ideal similarity measure will have an optimum value when the two images are registered. It will diverge from the optimum value for movements in any direction within the search space. It should diverge smoothly and rapidly and possess no local optima.

3.1.3 Search space

The search space determines the types of coordinate transformation which are allowed to register the images. For example if it is believed that a translation is required to align the images, then a sensible search space would consist of a set of all possible translations. As the number of degrees of freedom included in the search space increases so does the dimensionality of the search space, which causes the size of the search space to increase in an exponential fashion. Therefore, to register quickly it is important to keep the search space to the minimum number of dimensions required to achieve an accurate registration.

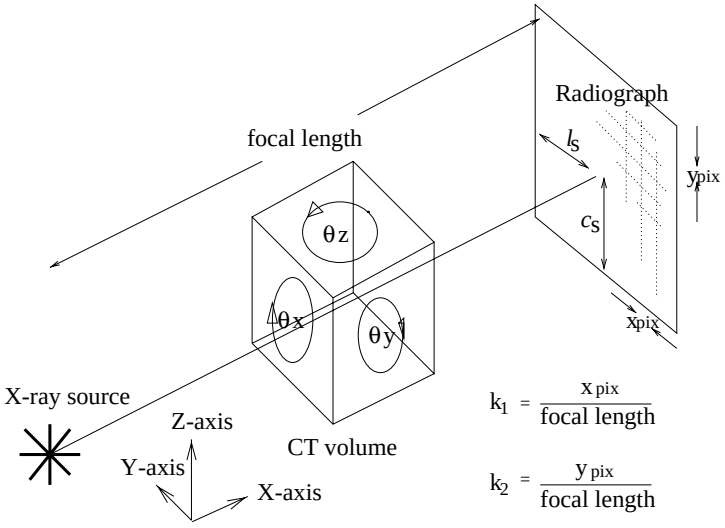


Figure 3.1: The ten degrees of freedom in rigid body perspective projection ($\mathbf{W} = (\theta_x, \theta_y, \theta_z, X, Y, Z, c_s, l_s, k_1, k_2)$). The extrinsic parameters, $\mathbf{E} = (\theta_x, \theta_y, \theta_z, X, Y, Z)$, define the position and orientation of the object being imaged with respect to the imaging device. The intrinsic parameters, $\mathbf{I} = (c_s, l_s, k_1, k_2)$, are associated with perspective projection: where c_s and l_s define the position where the normal to the imaging plane goes through the x-ray source, and k_1 and k_2 equal the pixel sizes in the horizontal (x_{pix}) and vertical (y_{pix}) directions respectively divided by the focal length

In 2D-3D image registration the search space can include all the transformations which may be present in 3D-3D matching i.e. translations, rotations, scaling, skew and deformation. Most of the current medical image registration algorithms assume a rigid body transformation and so have a search space which includes six degrees of freedom, although some of these algorithms have been extended to include scaling, skew [59, 117] and deformation [103]. In 2D-3D registration there are four additional parameters which define the perspective projection. Therefore, the search space for 2D-3D registration described by the perspective projection of rigid 3D motion can include up to ten

parameters, $\mathbf{W} = (\theta_x, \theta_y, \theta_z, X, Y, Z, c_s, l_s, k_1, k_2,)$. These ten parameters are shown in figure 3.1. In keeping with camera calibration terminology these parameters have been split up into two sets. The first set are extrinsic parameters, $\mathbf{E} = (\theta_x, \theta_y, \theta_z, X, Y, Z)$, which govern the position and orientation of the object being imaged with respect to the imaging device. Where $\theta_x, \theta_y, \theta_z$ are Euler angles representing the orientation of the 3D volume and X, Y, Z define the position of the centre of the 3D volume. The second set are intrinsic parameters, $\mathbf{I} = (c_s, l_s, k_1, k_2)$, which define the perspective projection. Where c_s and l_s define the position of the ray piercing point. The ray piercing point is the position on the imaging plane where the normal to the imaging plane goes through the x-ray source. The parameters k_1 and k_2 equal the pixel sizes in the horizontal and vertical directions respectively divided by the focal length.

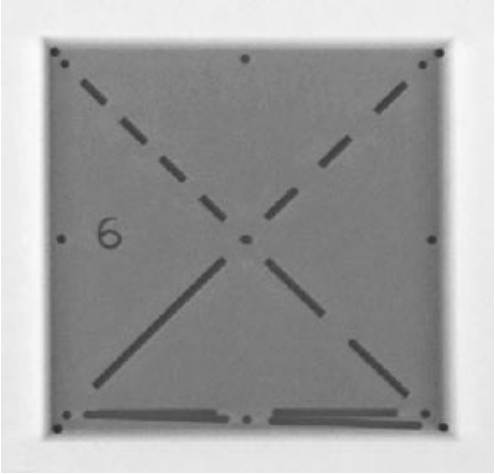


Figure 3.2: Fluoroscopy image of a calibration phantom. The phantom is an accurately machined acrylic cube with 14 embedded ball bearings, one at each corner of the cube and one in the centre of each cube face. The lines and number 6 (or 9) have been designed so that the orientation of the cube can be easily determined.

Optimising all ten parameters can produce a complex search space requiring a large number of iterations to reach registration. Because of this, most 2D-3D registering techniques calculate the intrinsic parameters prior to registration and then carry out a six degree of freedom rigid body registration. The intrinsic parameters can be found in a number of ways. Physical measurements can be taken when the images are acquired e.g. the focal distance can be measured. Another method is to use a calibration phantom. A fluoroscopy image of a calibration phantom is shown in figure 3.2. This phantom consists of an accurately machined acrylic cube with 14 embedded ball bearings. The relative positions of the ball bearings are known and by locating the 2D position of

the ball bearings in the fluoroscopy image it is possible to calculate the intrinsic parameters of the fluoroscopy set using a method described in section 4.2.1.

Another difference between 2D-3D and 3D-3D registration is the magnitude of the effect on the information extracted from each image due to alterations in different parameters. In 3D-3D image registration a translation along any axis will result in the same displacement for each voxel. Also, if the volumes are cubic, a rotation around any axis (which goes through the centre of a volume) will result in a similar average voxel displacement. This is not true for 2D-3D registration using perspective projection. Translations along different axes in 3D can have very different effects on the position of 3D structures when they are projected into 2D. The most marked example is comparing an in-plane with an out-of-plane translation. An in-plane translation is defined as a translation which is parallel to the imaging plane, whereas an out-of-plane translation is normal to the imaging plane. A numerical example is shown in figure 3.3. If the object in figure 3.3 is translated 1mm in the direction of the X-axis (an out-of-plane translation) then the position where the object projects onto the radiograph (point B) only alters by 0.022mm. Whereas, if the object is translated by 1mm in the direction of the Y-axis (an in-plane translation) then the position of point B alters by 1.1mm. Therefore, the effect on the position of point B of an in-plane translation is approximately fifty times larger than the effect of an out-of-plane translation.

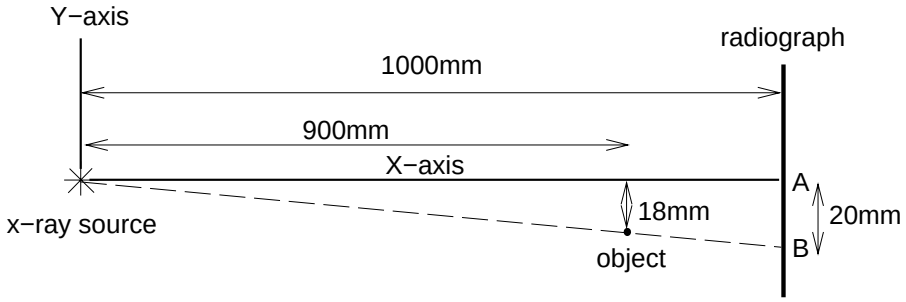


Figure 3.3: The difference in the position of a projected point caused by altering in-plane (along direction of Y-axis) and out-of-plane (along direction of X-axis) translational parameters. The projection of the object (point B) is initially 20mm from the position of the ray piercing point (point A). A 1mm in-plane translation will cause the position of point B to move fifty times further than a 1mm out-of-plane translation.

A similar effect can be observed when comparing in-plane and out-of-plane rotations. An in-plane rotation is defined to be about an axis which is normal to the imaging plane while out-of-plane rotations are about axes parallel to the imaging plane. The in-plane rotation, θ_x in figure 3.1, will on average have a larger effect on the position of projected points than the out-of-plane

rotations, θ_y and θ_z. One of the consequences of the difference between altering in-plane and out-of-plane parameters is that some degrees of freedom can be much more accurately constrained than others.

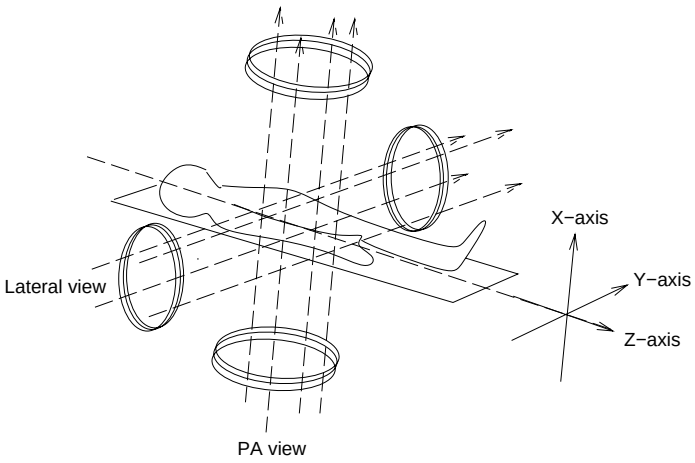


Figure 3.4: The in-plane degrees of freedom for PA and lateral fluoroscopy views. In the PA view translations along the Y-axis and Z-axis and rotations about the X-axis are in-plane. In the lateral view translations along the X-axis and Z-axis and rotations about the Y-axis are in-plane. Rotations about the Z-axis are always out-of-plane.

One solution to this difference between altering in-plane and out-of-plane parameters is to use multiple 2D images. If two 2D images are taken from different positions, then a parameter which was out-of-plane in one view may be in-plane (and, therefore, potentially easier to constrain) in the other view. If the two views are orthogonal then by using two 2D images the translational parameters can be well constrained, however, there will still be one rotational parameter which remains out-of-plane. For example, figure 3.4 defines a world coordinate system and shows the fluoroscopy gantry position for a posterior-anterior (PA) and a lateral fluoroscopy view. In the PA view translations along the Y-axis and Z-axis and rotations about the X-axis are in-plane. In the lateral view translations along the X-axis and Z-axis and rotations about the Y-axis are in-plane. Therefore, rotations about the Z-axis are always out-of-plane. Note that this is the same axis about which the fluoroscopy gantry rotated.

By using multiple views you can reduce the number of parameters which are always out-of-plane and so improve registration accuracy. However, to make use of multiple views the transformation between the views must be accurately known. This is possible using radiotherapy equipment, as accurate patient and beam positioning is already part of the clinical protocol. However, this is not the case when registering during an interventional procedure. There are a number of ways in which the transformation between the two views could be calculated, e.g. fiducial markers in the fluoroscopy field of view or by tracking the position of the fluoroscopy gantry. But

these methods would increase the complexity and reduce the flexibility of the fluoroscopy system. Therefore, a registration algorithm which requires only a single 2D image may be more suitable for use during interventional procedures.

3.1.4 Search strategy

The search strategy determines the best way to move through the search space in order to reach an optimum value of the similarity measure. Each calculation of the similarity measure has a computational cost and so a good search strategy will reach the optimal value of the similarity measure in the minimum number of iterations. The search strategy chosen depends greatly on the smoothness of the similarity measure within the search space.

There are a number of standard techniques for optimising functions [100]. Usually there is a trade off between the robustness of a search strategy and the number of iterations required to reach registration, where more robust strategies tend to require more iterations. The choice of search strategy depends on how well behaved the similarity measure is within the search space, where a plot of a well behaved similarity measure is assumed to form a smooth noise free shape with a single optimum position. A well behaved measure can use a strategy which requires a small number of iterations, while a badly behaved measure may require a more robust search strategy. The capture range for the registration algorithm is dependent on both the behaviour of the similarity measure within the search space and the choice of search strategy.

There are a number of differences between 2D-3D and 3D-3D registrations which should be taken into account when choosing a search strategy. Firstly, because altering out-of-plane parameters can have a much smaller effect on the position of features in 2D compared to altering in-plane parameters, the algorithm can be more susceptible to finding local minima. This is discussed in detail in section 5.4. Secondly, the search space may include four extra degrees of freedom, though, as mentioned earlier, the intrinsic parameters are usually calculated prior to registration and do not form part of the search space. Thirdly, the extra computational cost involved in transforming information either from 3D to 2D or vice-versa increases the importance of an efficient search strategy.

3.2 Feature based registration methods

This section provides a critical analysis of five feature based algorithms. Three of these algorithms [8, 51, 70] register segmented curves to surfaces. The earliest of these is the work by Lavallée *et*

al. [70, 71] which was first published in 1991 followed in 1995 by Betting and Feldmar [8, 43] and Guéziec *et al.* [51] in 1998. Also included in this section are the papers by Murphy [94] and Liu *et al.* [77], both of which were published in 1997. Both of these papers form DRRs from the CT volume from which they extract 2D features. A summary of the algorithms is included at the end of this section (see table 3.1).

3.2.1 Algorithm by Lavallée *et al.* [70, 71]

This paper describes an algorithm to match 3D smooth surfaces to two 2D contours. The algorithm uses as input a segmented 3D surface and two 2D segmented contours, as shown in fig. 3.5. The intrinsic parameters have been found previously using a *N-Planes BiCubic Spline Method* [19]. These intrinsic parameters are used to draw a three dimensional line from the 2D contour to the x-ray source. The minimum Euclidean distance between this line and the 3D surface is then calculated, d in figure 3.5. Distances within the 3D surface are taken to be negative, therefore, a distance of zero is only achieved if a 3D line is tangent to the surface. A number of points are picked around the 2D contours, lines drawn and minimum distances calculated. The overall similarity measure, S, is the sum of the squared minimum distances, see equation (3.1). A least squares minimisation of this measure is carried out using the Levenberg-Marquardt algorithm [100].

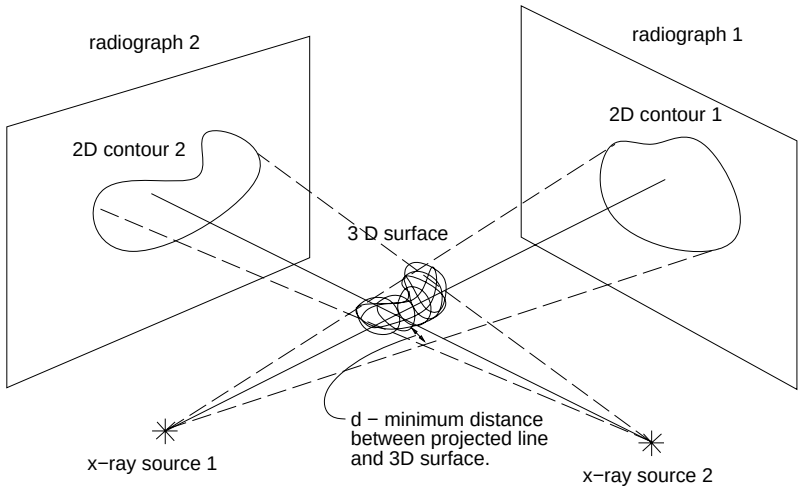


Figure 3.5: The use of projection lines to register a surface to two 2D contours.

$$S = \sum_{contour} \mathrm{d}^2 \tag{3.1}$$

The algorithm has been tested on simulated and anatomical 3D datasets where 2D contours were simulated by projecting the 3D surface. Because the 2D contours were simulated the datasets had a known "gold-standard" registration position. Using this data the algorithm registered robustly, always converged to the correct solution from a starting position 20^o away from the "gold-standard". The final registrations were accurate, with errors less than 1^o and 1mm.

An experiment was carried out using a single isolated vertebra which had two 3mm holes drilled through it. A CT scan was acquired of the vertebra, from which a 3D surface (made from 200,000 points) was segmented. The 2D images were produced using calibrated video cameras and an edge detection algorithm was employed to extract between 10 and 200 contour points. The algorithm registered the segmented 3D surface to the 2D contour points. The final registration parameters were used to guide a robotic arm, which held a laser, so that the laser beam passed down one of the holes in the vertebra. In all their experiments the alignment was visually perfect and the time taken from moving the vertebra to alignment of the laser beam was between 1 and 4 seconds.

A further experiment was carried out using a CT scan and two radiographs of a skull phantom. The skull was segmented from CT to give a 3D surface and the 2D contours were manually segmented from the radiographs. Four small catheters were placed into the skull prior to image acquisition and their centroids detected in the 2D and 3D images. After registration the 2D position of the catheters in the radiographs were backprojected to calculate their 3D position within the CT scan. The distance between the backprojected position of the catheters and the actual 3D position of the catheters was calculated. The final mean residual error was 1.7mm.

3.2.2 Algorithm by Betting and Feldmar [8, 42]

This paper describes an algorithm to register a 3D surface to a 2D contour. The algorithm uses only a single 2D image. The search space for the algorithm was defined by the six extrinsic parameters, though for one experiment the search space includes both the intrinsic and extrinsic parameters.

The algorithm begins with a segmented 3D surface and a segmented 2D contour. A set of 2D and 3D points are obtained from these segmented features using the fundamental properties of free form curves and surfaces. An example of a fundamental property relating a 3D line with its 2D projection is shown in figure 3.6. Both the position of these points and the direction of the normals to the 2D and 3D lines are calculated. The similarity measure calculates a four dimensional distance $d(m, M)$, between the 3D and 2D points, see equation (3.2), where M represents a 3D point position and m represents a 2D point position. Two of the dimensions in $d(m, M)$ relate to

the Euclidean distance between the 2D point position, m, and the position of the 3D point projected into 2D, $Proj(M)$. The other two dimensions relate to the difference between the direction of the normal to the 2D contour, $Normal_{2D}(m)$, and the projection of the 3D normal into 2D, $Normal_{3D}(M)$.

$$d(m, M) = (\beta(m - Proj(M))^2 + (Normal_{2D}(m) - Normal_{3D}(M))^2)^{1/2} \tag{3.2}$$

where β is a normalisation coefficient related to the 2D contour diameter.

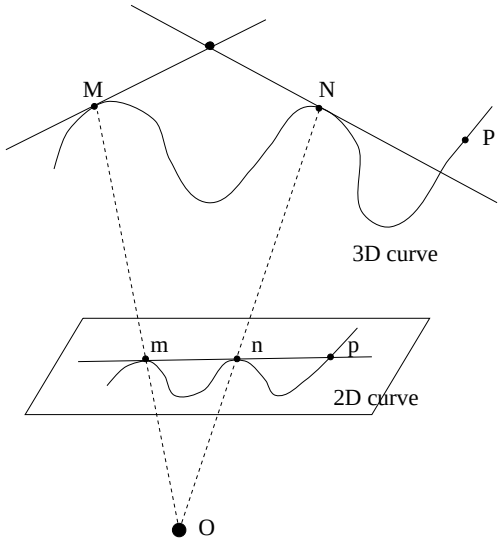


Figure 3.6: A fundamental property relating a 3D curve with its 2D projection. Two 2D points m and n which share the same tangent line are projections of two 3D points M and N such that the tangent lines at these points lie in the same plane. Moreover, if p is an intersection point between the tangent line and the 2D curve, then it is the projection of a 3D point P which is an intersection between the 3D curve and the plane which contains the tangent lines at points M and N. Reproduced from Betting and Feldmar [8].

A two-step algorithm is used, which is an extension of the Iterative Closest Point (ICP) algorithm of Besl and McKay [7]. Corresponding 2D and 3D points are linked together using a matching function, C, which attempts to minimise the similarity measure S (see equation (3.3)) while the rotational, $\mathbf{R}$, and translational, $\mathbf{t}$, parameters remain fixed. The matching function C finds a corresponding 3D point M for every 2D point m. The next step is to minimise S with respect to $\mathbf{R}$ and $\mathbf{t}$ with the correspondences fixed. The algorithm iterates around these two steps to achieve registration.

$$S(\mathbf{R}, \mathbf{t}, C) = \sum_{m_i} d(m_i, \mathbf{R}C(m_i) + \mathbf{t})^2 \tag{3.3}$$

The algorithm is used to register a CT scan to a video image of a mannequin head. The final registration error is given in terms of the final value of the similarity measure $d(m, M)$, i.e. the average Euclidean distance between 2D points and 3D points projected into 2D and the average angular distance between normals to the 3D lines projected into 2D and normals to the 2D lines. The average spatial error before registration was 2.1 pixels and the average angular error was 2 degrees. After registration these had fallen to 0.76 pixels and 0.17 degrees.

The algorithm was also used to register a CT image to a radiograph of a skull phantom. Both the extrinsic and the intrinsic parameters were optimised during the registration procedure. The algorithm registered from an initial error of 1.5 pixels and 1.9 degrees to a final error of 0.79 pixels and 0.7 degrees. Only one registration was reported using each dataset.

3.2.3 Algorithm by Murphy [94]

This algorithm is used to register fluoroscopy images to a CT volume to calculate the position of a patient during radiotherapy. A radiotherapy room has been designed which consists of a linear accelerator positioned by a robot arm, a treatment couch and two fluoroscopy sets. The aim of the system is to acquire biplanar fluoroscopy images both before and during the radiotherapy treatment and to register these images to the pre-operative CT scan. This registration provides information on the current patient position which is used to correctly position the linear accelerator as per the pre-operative treatment plan.

The algorithm produces DRRs from the CT volume from which 2D points on the edge of the skull are segmented. Corresponding points on the edge of the skull are segmented from the fluoroscopy images. To speed up the algorithm the DRRs are only produced in a number of small regions of interest which are likely to contain the edge of the skull. These regions of interest are established prior to treatment. The 2D point positions are determined by drawing lines, from the centre of the DRRs and from the centre of the fluoroscopy images, through the centre of the regions of interest. A Sobel edge detector [48] is used to calculate the position of the skull boundary along these lines. They also calculate the first order moment of the grey scale pixel values along these lines as an additional feature on which to register. The similarity measure used is the chi-squared statistic, equation (3.4), where A_i and B_i represent the information extracted from each image. Equal weights, w_i, are given to each component of the feature vector. Note: if the first moment of the grey scale pixel value is disregarded from the feature vector, then the similarity measure is simply the sum of the squares of the Euclidean distance between the 2D skull boundary points.

$$\chi^2 = \sum_i \frac{(A_i - B_i)^2}{w_i^2} \tag{3.4}$$

The search space was defined by the six extrinsic parameters and the search strategy begins by using a gradient search and switches to a Taylor series approximation as the optimum value is approached.

An experiment was carried out using a CT scan of a head phantom from which DRRs were produced to form simulated 2D datasets. Since the 2D images were simulated, the "gold-standard" position was known. This position was perturbed by up to 8mm and 5° along each of the extrinsic parameters and the algorithm was used to register the CT volume to the DRRs. Twenty registrations were carried out and the average of the final root mean square (RMS) errors was 0.3mm for the translational parameters and 0.63° for the rotational parameters.

An experiment was also carried out to investigate the ability of the algorithm to track a patient's position if they move during a radiotherapy procedure. The same head phantom was used, though this time the simulated 2D images were replaced by fluoroscopy images. The movement of the head was either measured mechanically or by using an electronic tilt meter. The head phantom was moved by up to 3mm or degrees, then new fluoroscopy images were acquired and the algorithm was used to register to the new position. The RMS error ranged between 0.5 to 1.0 mm and 0.6 to 1.3° depending on the direction of movement. The time taken to register was approximately 2 seconds.

3.2.4 Algorithm by Liu *et al.* [77]

This algorithm was used to register a CT scan to two radiographs of a dry skull. The algorithm produces DRRs from the CT scan. A number of curves are extracted from the two DRRs and from the two radiographs using a core based approach [79]. These 2D curves are back projected into 3D to form a set of 3D space curves termed "synthesised anatomical structures". The curves are defined as synthetic because they are formed from 2D structures which may not have a corresponding 3D counterpart. The registration is based on these 3D "synthesised anatomical structures" extracted from the DRRs and the radiographs.

A registration was carried using the Visible Human CT dataset where the 2D images were simulated by producing DRRs. Because the 2D data was simulated the "gold-standard" registration is known. The final registration error was less than 1mm over a large portion of the 2D field of view.

Two further registrations were carried out, this time using a skull phantom. Two CT scans

were taken of the phantom, one with a cannula inserted and one without. The first registration again used synthetic 2D images. Two DRRs were created using the CT dataset with the cannula inserted. These DRRs were used to represent intra-operative fluoroscopy images. The CT scan without the cannula inserted was then registered to these DRRs. The final registration parameters were used to back project the 2D positions of the cannula (in the DRRs) to calculate a 3D location of the cannula within the CT volume. This 3D position was then compared with the actual 3D position of the cannula and the registration error was calculated to be 1.6mm. The above experiment was also repeated using radiographs for the 2D images and the final error was 2.1mm.

3.2.5 Algorithm by Guéziec *et. al* [51]

This algorithm is used to register a CT scan to fluoroscopy images of a cadaver bone. The algorithm uses a semi-automated deformable model technique [63] to detect the 2D bone contours in both the fluoroscopy image and in each slice of the CT volume. The CT bone contours are used to construct a 3D surface which consists of a set of triangles. The algorithm determines, for a particular set of intrinsic and extrinsic parameters, which triangles are facing towards and which triangles face away from the x-ray source. The boundary between these two sets of triangles defines the apparent contour of the bone when projected into 2D. The similarity measure employed by this algorithm is similar to that used by Lavallée *et al.* [70, 71]. Lines are drawn from the x-ray source to the 2D contour and the algorithm attempts to minimise the minimum distance between these lines and the apparent contour of the bone.

The clinical motivation behind this work was to guide a surgical robot used in hip replacement procedures. The current registration method uses external fiducial markers drilled into the femur prior to imaging. This increases patient discomfort and requires a separate procedure, therefore, a less invasive registration method would be preferred. The experiment described in the paper was carried out on a cadaver femur and care was taken to simulate clinical conditions. A "gold-standard" registration was calculated using the current clinical protocol of attaching fiducial marker pins to the femur. The "gold-standard" position was perturbed by up to 7 degrees and the algorithm used to recover the parameters. Eleven registrations were carried out and the final registration error measured at each marker location ranged from 1.2 to 4.8mm.

3.2.6 Combined segmentation and registration

The processes of segmentation and registration can be mutually beneficial. Feature based algorithms have been shown to register quickly, accurately and robustly when provided with accurately

Algorithm authors	Lavallée *et al.* 1995	Betting & Feldmar 1995	Murphy 1997	Liu *et al.* 1997	Guéziec *et al.* 1998
Information extracted	3D surfaces 2D contours	3D & 2D points	2D edges	synthesized anatomical structures	3Dsurfaces 2D contours
Similarity metric	distance between projection lines and a surface.	2D distances between points	chi-squared statistic	not stated	distance between projection lines and 3D contour.
Search space	extrinsic	extrinsic + intrinsic	extrinsic	extrinsic	extrinsic
Search strategy	Levenberg-Marquardt	iterative closest point	gradient search and Taylor approx.	not stated	robust M-estimator
Speed of registration	1 - 5secs	10-20 secs	2 secs	10 mins	1 sec
Computer used	DEC station 5000/200	DECalpha	200MHz machnine	not stated	IBM RS6000 580
Images used	simulated, video, phantom CT and radiographs	video, CT and radiograph of phantom.	simulated, phantom CT and radiographs	simulated and phantom CT and radiographs	CT and fluoroscopy of a femur
Segmentation method	video - edge extraction. x-ray - manual	not stated	Sobel edge detector	core extraction with manual interaction.	semi automatic, using active contour models.
Number of views	2	1	2	2	2
Method of determining accuracy	synthetic data and fiducial point	final value of similarity measure	synthetic data and a tilt meter	fiducial point	fiducial markers
Error measured in 2D or 3D	3D	2D	3D	3D	3D
Registration accuracy	0.2mm, 0.16° synthetic data, 1.7mm phantom	0.8 pixels	0.5 to 1.0mm, 0.6° to 1.3°	1mm synthetic data, 2mm phantom	1.2 - 4.8mm

Table 3.1: Summary of the state of the art 2D-3D feature based registration algorithms in the literature.

segmented data. Also, consider how much easier it would be to segment bony structures from a fluoroscopy image if information from a registered CT image (where bone is relatively easy to segment) could also be used. Hamadeh *et al.* [55] propose a unified approach, combining registration and segmentation based on bone surfaces in CT and contours in radiographs. The registration approach is very similar to Lavallée et al. [71, 70] though instead of manually segmenting the 2D contours they use the following method. The probability of a given pixel lying on a contour depends on a number of factors. Some of these factors are obtained from the radiograph, such as local gradient or the length of a connected component of contour pixels. However, the CT volume is also used to obtain information on the position of 2D contours. The bony outline from CT is projected onto the radiograph and the distance between this outline and a given pixel effects the probability of that pixel belonging to a bony contour. The paper does not report the results from any registrations but does present an interesting approach to solving the problems involved with segmentation when carrying out feature based 2D-3D registration.

3.3 Intensity based registration methods.

This section provides a critical analysis on three of the best intensity based 2D-3D registration algorithms. The first of these is by Lemieux *et al.* [73] which was published in 1994 and registers a CT volume to two radiographs of a head phantom. The next method is by Gilhuijs *et al.* [45] which was published in 1995 and was used to register portal images to a CT volume. Brown *et al.* [49] published an algorithm in 1996 which registered a CT volume to a single radiograph of a dry femur. These algorithms are summarised in table 3.2 which can be found at the end of this section.

3.3.1 Algorithm by Lemieux *et al.* [73]

This paper describes an algorithm to register a CT volume to an anterior-posterior (AP) and a lateral radiograph of a skull phantom. The search space was defined by the six extrinsic parameters. The intrinsic parameters were calculated using fiducial markers which were attached to the skull phantom while the CT volume and the radiographs were acquired. The algorithm produces DRRs from the CT volume and compares them to the radiographs using two similarity measures, linear cross correlation and a gradient product measure. The linear cross correlation coefficient (see [100] or section 5.2.1 for details) was calculated between the intensity values in the AP radiograph and in the AP DRR to give R_{AP}, and also between the lateral radiograph and the lateral DRR to give R_{lat}. These two values were combined to give R_{total} as shown by equation (3.5), where N_{AP} and N_{lat} are the number of overlapping pixels which were used to calculate R_{AP} and R_{lat} respectively.

$$R_{total} = -\frac{R_{AP}N_{AP} + R_{lat}N_{lat}}{N_{AP} + N_{lat}} \tag{3.5}$$

The gradient product measure calculates two dimensional gradients in the AP and lateral radiographs, $\mathbf{g}(AP)$ and $\mathbf{g}(lat)$ and also in the AP and lateral DRRs, $\mathbf{g}(AP\ DRR)$ and $\mathbf{g}(lat\ DRR)$. They are combined in the following manner to give the gradient product similarity measure, G.

$$G = -\frac{1}{N_{AP} + N_{lat}} \left(\sum_{i=1}^{N_{AP}} [\mathbf{g}(AP)_i \cdot \mathbf{g}(AP\ DRR)_i] + \sum_{i=1}^{N_{lat}} [\mathbf{g}(lat)_i \cdot \mathbf{g}(lat\ DRR)_i] \right) \tag{3.6}$$

The cross correlation measure was found to possess fewer local optima at a distance from the registration position, while the gradient product measure was found to be smoother close to the registration position. Because of this, the cross correlation measure was used in the early stages of the registration process and when close to the registration position the similarity measure switches to the gradient product measure.

The search strategy was a two stage, multi-resolution process. The first stage was orientation initialisation, when the cross correlation similarity measure was used. This stage optimised five of the six extrinsic parameters, all except θz which was assumed to be near to optimum, note that θz is the degree of freedom which is out-of-plane in both radiographs, see figure 3.4. When orientation initialisation was complete a full 3D registration based on Powell's multidimension set method [100] was carried out using the gradient based similarity measure.

A "gold-standard" registration was found using fiducial markers. A start position for the algorithm was found by randomly altering the extrinsic parameters by $\pm$ 10 mm or degrees from the "gold-standard" position, for all the parameters except θz, which is only altered by $\pm 5^{\circ}$. One hundred registrations were carried out. The registration error was calculated by taking a number of points distributed about the centre of mass of the skull. These points were transformed using the "gold-standard" matrix and also by the final registration matrix. The mean 3D distance between these two sets of points was reported as the registration accuracy. The results showed a high degree of accuracy. In only one of the one hundred registrations was the mean distance larger than 1mm. The time taken to achieve registration was half an hour. The algorithm has been used clinically to verify the placement of intracerebral electrodes by registering post-operative radiographs with pre-operative CT data.

3.3.2 Algorithm by Gilhuijs *et al.* [45]

This algorithm has been designed to register a CT volume to portal images. The algorithm locates prominent bone ridges in the portal images using either a top-hat transformation or multiscale medial axis enhancement [79]. Rays are cast from the x-ray source to these ridge lines and the Hounsfield values are integrated along the ray path. The hypothesis is that bone ridges will be located where there is a maximum path traversed through the bone. Therefore, when the sum of these integrals is a maximum, registration has been achieved. By reducing the number of rays cast through the CT volume the speed of the algorithm is increased. I was unsure whether this algorithm falls into the feature based or intensity based category, as features have been located in the portal images. I opted to include this algorithm in the intensity based section because, although bone ridges are extracted from the 2D images, the actual similarity measure is based on the integral of voxel intensities.

The search space for the algorithm was defined by the extrinsic parameters. The intrinsic parameters were calculated using information from the set up of the radiotherapy equipment (e.g. gantry and table settings) and also the position of field-defining wires and central-cross wires in

the portal images. The algorithm was tested on portal images from six patients where rotational misalignment was suspected. The algorithm produced a good visual match and the results were comparable, within 1^o and 1mm, to a manual registration method. Also, a cross validation type of approach was used, where the registration results obtained with one pair of portal images were compared with the results from another pair of portal images acquired during the same treatment session. The two sets of of images produced comparable registration results which were approximately within 1^o and 1mm of each other. The algorithm took less than two minutes to reach registration.

3.3.3 Algorithm by Brown *et al.* [49]

This algorithm was used to register a CT volume to a radiograph of a dry femur. The algorithm produces DRRs from the CT volume which are compared to the radiograph using a gradient correlation similarity measure. This measure is very similar to the measure described in section 5.2.4. Gradient images are formed from the radiograph and from the DRRs using a 3x3 Sobel filter [48]. Two images are formed from the radiograph and from the DRR, these two images represent the vertical and horizontal gradients in each image. Normalised cross correlation [100] is carried out between the vertical gradients in the radiograph and in the DRR and also between the horizontal gradients in the radiograph and in the DRR. The final value of the similarity measure is the average of these two correlation values.

The search space for the algorithm was defined by the six extrinsic parameters. Experiments were carried out using a dry femur bone which had been attached to a precisely machined calibration object. This calibration object was used to calculated the intrinsic parameters and a "gold-standard" registration. Powell's multidimensional set method [100] was used to optimise the similarity measure.

An experiment was carried out using simulated 2D images rather than a radiographs. Each of the "gold-standard" extrinsic parameters were altered by up to $\pm$ 1mm or degree and five registrations were carried out. Registration accuracy was measured by randomly picking 100 points within a 100mm cube centred on the femur. These points were transformed by the "gold-standard" registration matrix and also by the registration matrix calculated using the registration algorithm. The average 3D distance between these two sets of points was taken to be the registration accuracy. The final registration accuracy was 0.96mm. This experiment was then repeated a further two times when the size of the alteration increased from $\pm$ 1mm or degree to $\pm$ 2 and then to $\pm$ 3 mm or degrees. The final registration accuracy was 1.9mm and 5.2mm respectively.

Algorithm authors	Lemieux *et al.* 1994	Gilhuijs *et al.* 1996	Brown *et al.* 1996
Information extracted	Radiographs and DRRs	Portal images and DRRs	Radiographs and DRRs
Similarity metric	cross correlation, Gradient correlation	maximum integral of Hounsfield values	Gradient correlation
Search space	extrinsic	extrinsic	extrinsic
Search strategy	Powells	downhill simplex	Powells
Speed of registration	30 mins	2 mins	not stated
Computer Used	SUN SPARCstation 10/30	90MHz Pentium PC	not stated
Images used	radiographs and CT of skull phantom.	portal images 2D, and CT from clinical cases	radiographs and CT of dry femur
Segmentation Method	none	top hat transformation on portal image	none
Number of views	2	2	1
Method of determining accuracy	fiducial markers	multiple views. cf. to manual registration	fiducial markers
Error measured in 2D or 3D	3D	3D	3D
Registration Accuracy	1mm	1 ^{o}, 1mm	1mm case 1, 3mm case 2

Table 3.2: Summary of the state of the art 2D-3D intensity based registration algorithms in the literature.

An experiment was also carried out using radiographs. The "gold-standard" parameters were perturbed by $\pm$ 2 mm and degrees to obtain a starting position for the algorithm. Registrations were carried out using two datasets and registration accuracy was calculated as in the previous experiment. The final registration error was less than 1mm for the first dataset and nearly 3mm for the second dataset. If the extrinsic parameters were perturbed by more than $\pm$ 4 mm or degrees the algorithm typically did not return to the "gold-standard" position.

3.4 Registration of CT volumes to x-ray projection images

The aim of this section is to discuss the performance of the algorithms when registering to x-ray projection images, as opposed to simulated images. A summary of the performance of the algorithms when registering to x-ray projection images is shown in table 3.3. This table includes a description of the imaged object, the number of registrations reported, the accuracy of these registrations and the independent registration method used to calculate the registration accuracy.

The table shows that seven of the eight algorithms were tested by registering phantom images. Only the algorithm by Gilhuijs *et al.* [45] was tested using clinical images. Overall, the number

Algorithm	Imaged object	No. of registrations	Accuracy	Independent accuracy measure
Lavallée *et al.*	skull phantom	1	1.7mm	fiducials
Betting and Feldmar	skull phantom	1	0.8 pixels	none
Murphy	skull phantom	86 (approx)	0.5-1mm, 0.6-1.3^o	tilt meter
Liu *et al.*	skull phantom	1	2mm	fiducial
Guéziec *et al.*	cadaver femur	11	1.2-4.8mm	fiducials
Lemieux *et al.*	skull phantom	300	1mm	fiducials
Gilhuijs *et al.*	clinical, neck and pelvis	6	1mm, 1^o	manual registration & cross validation
Brown *et al.*	femur	2	1-3mm	fiducials

Table 3.3: Summary of algorithms when registering CT volumes to x-ray projection images. The double horizontal line separates the feature based algorithms from the intensity based algorithms.

of registrations to x-ray projection images was low. Only the algorithms by Guéziec *et al.* [52], Murphy [94] and Lemieux *et al.* [73] carried out more that ten registrations to x-ray projection images. Also, three of the five feature based algorithms only carried out a single registration to an x-ray projection image. One of the benefits of using a feature based method is that they can register quickly, therefore, it is surprising that only a single registration was carried out using some of these algorithms. The reported registration accuracy ranged from submillimeter to 4.8mm. Betting and Feldmar [8] quote the final value of their similarity measure as registration error. This is not an independent measure of registration accuracy and so should be treated with caution.

For some of the algorithms the size of the final registration error was fairly large when it is compared with the distance between the starting estimate for registration and the "gold-standard" position. For example, the registration algorithm designed by Murphy [94] reported a final registration error of 1.3 degrees when the starting position was only 3mm or degrees away from the registration position in one or two degrees of freedom. Similarly Brown *et al.* [49] report errors of 3mm when the starting position is $\pm$ 2mm or degrees from the "gold-standard" position in all six extrinsic parameters.

Registration speed was not compared in table 3.3 because some of papers gave insufficient detail to make a meaningful comparison possible. The speed of registration will depend on a number of factors such as the speed of the computer used and the size of the images (or the number of features extracted) on which the algorithm was tested. Also, a number of the feature based algorithms do not state the time taken to carry out the feature extraction process.

Figure 3.7 shows the relationship between 3D surface features in a CT image and 2D lines in a fluoroscopy image of a vertebra. Only the vertebral body and the transverse processes are

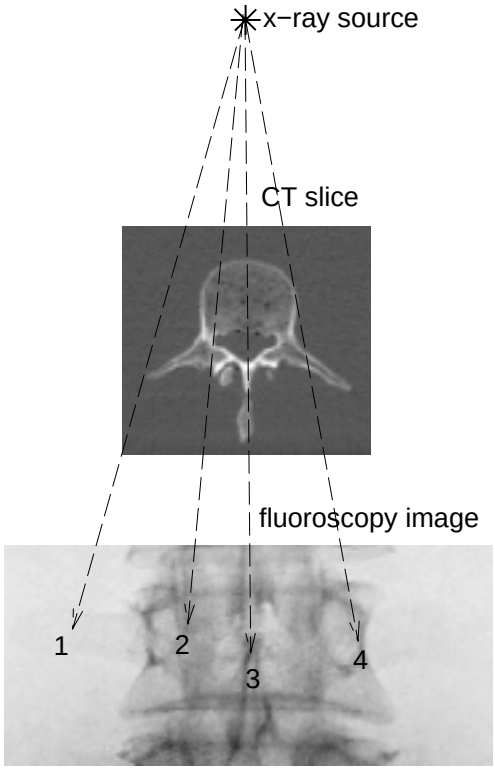


Figure 3.7: The relationship between 2D fluoroscopy lines and 3D CT surfaces for a lumbar vertebrae. Features shown are 1. transverse process, 2. pedicle, 3.spinous process and 4. vertebral body.

responsible for the external contours in the fluoroscopy image. The lines in the 2D image created by the pedicles and spinous process are caused by x-rays which travel through the vertebral body. Because the transverse processes are often not visible in clinical AP fluoroscopy images and because the vertebral body is fairly symmetric, most of the registration information comes from registering on the pedicles and spinous process. Therefore, an algorithm which only registers on the external contours from the 2D image, such as the algorithm by Lavallée *et al.*, would be severely disadvantaged when attempting to register the CT and the fluoroscopy images shown in figure 3.7.

3.5 Feature based vs Intensity based methods

I have chosen to design an intensity based 2D-3D registration algorithm in this thesis. My reasons are as follows. Firstly, intensity based 2D-3D registration algorithms appear to be more accurate than feature based algorithms when registering to x-ray projection images. This is shown in table

3.3 by two of the three intensity based algorithms registering to an accuracy of 1mm or less, whereas only one feature based algorithm, the algorithm by Murphy [94], achieved this accuracy.

Secondly, most intensity based methods do not require a feature extraction step. Automatic fast and accurate feature extraction from a complex scene such as an interventional fluoroscopy image is a difficult task [70]. Also, the paper by Guéziec *et al.* [51] describes experiments using synthetic data which indicated that segmentation was the foremost source of their errors.

Thirdly, 3D-3D multi-modality intensity based image registration algorithms [27, 113] have been shown to produce very accurate registrations [137]. Also, in West *et al.* [136] the performance of intensity based and surface based 3D-3D multi-modal image registration algorithms were compared. The results indicated that intensity based techniques gave substantially more accurate and reliable results than surface based techniques. I hypothesise that a similar situation will occur in 2D-3D image registration and that intensity based algorithms will be shown to be more accurate and robust than their feature based counterparts. Please note, I am primarily talking about registrations of CT volumes to radiographs or fluoroscopy images. Another example of 2D-3D registration is the registration of CT angiograms to DSA images. In this case, because it is generally easier to segment vessels from a DSA image than bony outlines from a fluoroscopy image, the above arguments for choosing an intensity based approach may be less valid.

3.6 Registering on single or multiple views

The majority of the algorithms reviewed in this chapter require two or more 2D images to register to the 3D volume. The exceptions are the algorithms described in Betting and Feldmar [8, 43] and Brown *et al.* [49]. If multiple 2D images are used then the transformation between the images must be known. This can be achieved by two methods. Firstly, the imaging device can be tracked either mechanically or using an optical tracking device. This method is particularly useful for radiotherapy applications where the gantry has been designed so that accurate measurements of its position and orientation can be obtained. The second method is to use a calibration object (e.g. the cube shown in figure 3.2). If this object remains in a rigid body relationship with the patient and it is present in both the 2D images, then the transformation between the two fluoroscopy views can be calculated. Neither of these methods are ideally suited for use during an interventional procedure using fluoroscopic guidance. A typical fluoroscopy gantry has not been designed so that it can be accurately positioned. Also, the requirement for accurate positioning may effect the flexiblilty of fluoroscopy as an imaging device. Optical tracking is an option, but it is expensive. A

calibration object could be used during interventional procedures, though problems may arise as the range of fluoroscopy views would be limited to those which keep the calibration object in the field of view. Algorithms which only require a single fluoroscopy image do not have any of these restrictions and can be used to register images which have been acquired using current clinical protocols.

3.7 Conclusions

To conclude, although a number of 2D-3D algorithms have been designed, the majority of these algorithms have only been applied to register images of a phantom skull or femur. I intend to design a registration algorithm to a register a clinical fluoroscopy image to a pre-operative CT volume of the abdominal area for use in image guided interventions. This has not been previously attempted. I believe that an intensity based 2D-3D registration algorithm registering to a single fluoroscopy image is the best choice of algorithm for my purposes.

Chapter 4

Producing digitally reconstructed radiographs

This section describes the information extraction step for the registration algorithm. It is titled "Producing digitally reconstructed radiographs" as most of the work carried out in the information extraction step is involved with the production of DRRs from the CT volume. However, in order to compare information from the two modalities it is also advantageous to apply some image processing steps to the fluoroscopy image; these are dealt with in section 4.4. The section begins with an overview of the imaging modalities, which is followed by a discussion on the mathematics involved with perspective projection. Then the method to produce DRRs is detailed, followed by how the fluoroscopy images are altered to more closely resemble DRRs.

4.1 Overview of the imaging modalities

The registration algorithm is designed to register a fluoroscopy image to a CT volume. As discussed in section 3.1.1 it is important to take into account the underlying physical principles behind each of these modalities, so that compatible information can be extracted from each image. The next two sections outline the physical principles behind firstly, the production of fluoroscopy images and secondly, the production of CT images. They are brief outlines, which concentrate on the differences between the modalities which will effect the registration algorithm.

4.1.1 Production of fluoroscopy images

Fluoroscopy images are transmission images created by a beam of x-rays. These rays are attenuated by varying amounts as they pass through the patient. The amount of attenuation depends

on the physical characteristics of the materials which they pass through. Therefore, the emergent beam contains information on the internal structure of the patient. This emergent beam hits a phosphor screen which converts the x-ray photons into light photons. In early fluoroscopy systems the phosphor screen was viewed directly by a radiologist in a darkened room. The advent of the image intensifier allowed the radiation dose to the patient and radiologist to be reduced and removed the requirement for a darkened room. Image intensifiers use a photocathode to convert light photons from the phosphor screen into electrons. The electrons are accelerated in a magnetic field (to intensify the image) and converted back to light photons using an output phosphor screen. The image can then be stored digitally using a CCD camera. These cameras typically produce images of size 512×512 or 1024×1024 pixels. At high magnification the physical size of each pixel, at the input to the image intensifier, can be very small (e.g. 0.17×0.17mm on a Philips MD3 fluoroscopy set).

Image intensifiers can cause two types of geometric distortion. The first type is caused by the mapping of a "flat" image onto a curved input phosphor. This causes pincushion distortion. The second type is caused by stray magnetic fields, such as those from the earth, interacting with the electrons in the image intensifier which causes s-curve distortion. Pincushion distortion is independent of the position of the image intensifier, while s-curve distortion depends on the angle of the image intensifier with respect to any stray magnetic fields. A considerable change in the magnitude and shape of the s-curve distortion pattern can be observed as the orientation of the image intensifier is altered [40]. The amount of distortion varies between image intensifiers, it is usually largest at the edge of the images, where it can be as large as 10mm [106].

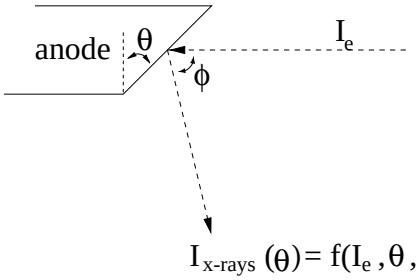


Figure 4.1: The intensity of x-rays emitted from the anode (I_{x-rays}) is a function of the incident intensity of electrons (I_e) and also the angles between the incident electrons and emitted x-ray (ϕ) and anode target angle (θ).

The x-ray beam is generated using a conventional x-ray tube which is rigidly linked and permanently centred on the image intensifier using a c-arm. The x-rays are produced by a beam of energetic electrons, accelerated through a potential difference of thousands of volts, striking a

target (usually tungsten). The electrons are decelerated by interactions with target atoms and some of the energy they lose appears in the form of x-ray photons. The effective energy of these photons is typically between 50-60keV. The intensity of the x-ray beam depends not only on the intensity of the incident electrons, but also on the angles between the incident electrons, the emitted x-ray and the anode target angle, see figure 4.1; this is known as the "anode-heel" effect. The result of the "anode-heel" effect is a gradual shading between the bottom and the top of the fluoroscopy image. The image intensifier also has a non-uniform response, with images slightly brighter in the centre of the field of view and darker at the periphery.

Two fluoroscopy sets were used to acquire images for this dissertation; a Philips MD3 system, see figure 4.2, and a mobile Philips BV29 system. Both of these sets have a fixed focal distance (1125mm for MD3 and 1000mm for BV29) and three magnification settings.

Figure 4.2: A Philips Multi Diagnost 3 (MD3) fluoroscopy set in the Kay Glendinning Interventional Suite, Guy's Hospital.

4.1.2 Production of computed tomography images

The work of Hounsfield in 1973 [60] led to the development of the first commercial CT system. CT produces 2D images of slices through the body using a tomographic method. By imaging a number of adjacent slices a 3D image can be constructed. The voxel intensities in a CT scan are given in Hounsfield units H, as shown in equation (4.1), where μ represents the linear attenuation coefficient of the voxel and μ_w is the linear attenuation coefficient of water.

$$H = \frac{\mu - \mu_w}{\mu_w} 1000 \tag{4.1}$$

The CT scanner used to produce images for this dissertation was a Philips TOMOSCAN SR 7000, see figure 4.3. The scanner uses x-rays of an effective energy of around 80keV and can produce either spiral CT scans or a series of single slice images. The voxel dimensions of images from this machine are typically around 1mm in-slice with a slice thickness varying between 1 and 10mm. The slice thickness of diagnostic scans is very dependent on the clinical case, a major factor being the radiation dose to the patient. It is common to tilt the CT gantry to obtain a more clinically useful slice orientation. This results in an altered slice thickness and non-orthogonal, skewed, image axes.

Figure 4.3: Philips TOMOSCAN SR 7000 CT scanner at Guy's Hospital

4.2 Mathematics to achieve perspective projection

The mathematics to project a 3D point onto a 2D plane using a perspective projection transformation matrix is well known. A discussion of this topic is given in [38] and most camera calibration [102, 122] and perspective matrix decomposition [44, 111] papers include an overview on the mathematics involved. The perspective projection transformation used to produce the DRRs has ten degrees of freedom, $\mathbf{W} = \{\theta_x, \theta_y, \theta_z, X, Y, Z, c_s, l_s, k_1, k_2\}$; these are shown in fig. 3.1. These parameters can be used to create a 3×4 perspective projection matrix $\mathbf{M}$. This matrix can

be used to project a 3D position in homogeneous world coordinates $(x, y, z, 1)^T$ to a 2D position in homogeneous fluoroscopy coordinates $(c, l, 1)^T$, see equation (4.2).

$$\mathbf{M} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} c \\ l \\ 1 \end{pmatrix}, \tag{4.2}$$

where λ is a scaling factor. The matrix $\mathbf{M}$ can also be used to transform a 2D fluoroscopy position $(c, l, 1)^T$ to a 3D line in world coordinates $(x, y, z, 1)^T + s\hat{\mathbf{n}}$ where s is a scalar and $\hat{\mathbf{n}}$ is a 3D unit vector. Due to the scaling factor and the use of homogeneous coordinates it is the relative magnitudes of the elements in matrix $\mathbf{M}$ which are important, rather than their absolute values and so it is customary to divide through the matrix by the bottom right hand element [44].

The matrix $\mathbf{M}$ can be split up into three separate matrices, a 3×4 perspective matrix $\mathbf{P}$, a 4×4 translational matrix $\mathbf{T}$ and a 4×4 rotational matrix $\mathbf{R}$. Equation (4.3) shows how these matrices are combined to produce $\mathbf{M}$ and equations (4.4), (4.5) and (4.6) show how $\mathbf{P}$, $\mathbf{T}$ and $\mathbf{R}$ are formed from the intrinsic and extrinsic parameters. I have chosen to rotate the coordinate system before translating, as in Tsai[122] and Weng *et al.* [135]. This ordering decouples the rotational and translational parameters. If the translation is carried out first, as in Ganapathy [44] and Rougée *et al.* [102], then the initial translation is also rotated by matrix $\mathbf{R}$ and so a small change in the rotational parameters can also cause a large translation. The final matrix in both of these cases are equally valid, but it can be harder to relate the translational parameters to the world coordinate system in the second case.

$$\mathbf{M}(\mathbf{W}) = \mathbf{P}(c_s, l_s, k_1, k_2)\mathbf{T}(X, Y, Z)\mathbf{R}(\theta_x, \theta_y, \theta_z) \tag{4.3}$$

$$\mathbf{P} = \begin{pmatrix} c_s & 1/k_1 & 0 & 0 \\ l_s & 0 & 1/k_2 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{4.4}$$

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & X \\ 0 & 1 & 0 & Y \\ 0 & 0 & 1 & Z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

$$\mathbf{R} = \begin{pmatrix} \cos\theta_z \cos\theta_y & \sin\theta_z \cos\theta_x + \cos\theta_z \sin\theta_y \sin\theta_x & \sin\theta_z \sin\theta_x - \cos\theta_z \sin\theta_y \cos\theta_x & 0 \\ -\sin\theta_z \cos\theta_y & \cos\theta_z \cos\theta_x - \sin\theta_z \sin\theta_y \sin\theta_x & \cos\theta_z \sin\theta_x + \sin\theta_z \sin\theta_y \cos\theta_x & 0 \\ \sin\theta_y & -\cos\theta_y \sin\theta_x & \cos\theta_y \cos\theta_x & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4.6)$$

4.2.1 Calculating the perspective projection matrix

In this thesis the perspective projection matrix $\mathbf{M}$ has been calculated using two different methods. Firstly, the following method is used by the registration algorithm to calculate projection matrices so DRRs can be produced. This method uses the known values of the intrinsic parameters (c_s, l_s, k_1, k_2) and the current best estimate of the extrinsic parameters (θ_x, θ_y, θ_z, X, Y, Z) to calculate $\mathbf{M}$ using equations (4.3), (4.4), (4.5) and (4.6). The intrinsic parameters are accurately calculated prior to registration using a suitable calibration object. The extrinsic parameters are either at the starting position or at the current best registration position based on the value of a similarity measure.

The second method is used to calculate a "gold-standard" perspective projection matrix $\mathbf{M}_{GS}$ based on the position of fiducial markers. If a number, n, of 3D point positions, $(x_i, y_i, z_i, 1)^T, i = 1, \ldots, n$, and their 2D projections, $(c_i, l_i, 1)^T, i = 1, \ldots, n$ are known, then $\mathbf{M}_{GS}$ can be calculated by solving a set of simultaneous equations. For each pair of 2D and 3D points we can expand the matrix in equation (4.2) to obtain two linear equations (4.7) [48], where $m_{jk}, j = 1, \ldots 3, k = 1, \ldots 4$ represent the individual elements in matrix $\mathbf{M}_{GS}$. There are twelve unknowns in the 3×4 matrix $\mathbf{M}_{GS}$ and each pair of points provides two equations. Therefore, at least six points are required to calculate the matrix coefficients. More commonly a larger number of points are used and a singular value decomposition (SVD) method [100] is used to find the least squares solution to the system of linear equations.

$$\begin{aligned} m_{11}x_1 + m_{12}y_1 + m_{13}z_1 + m_{14} &= (m_{31}x_1 + m_{32}y_1 + m_{33}z_1 + m_{34})c_1 \\ m_{21}x_1 + m_{22}y_1 + m_{23}z_1 + m_{24} &= (m_{31}x_1 + m_{32}y_1 + m_{33}z_1 + m_{34})l_1 \end{aligned} \quad (4.7)$$

4.2.2 Decomposing the perspective projection matrix

To relate the perspective projection matrix $\mathbf{M}$ to the world reference frame it is useful to be able to decompose $\mathbf{M}$ into the perspective parameters $\mathbf{W}$. Ganapathy [44] and Strat [111] have both proposed methods to decompose $\mathbf{M}$ into $\mathbf{W}$. These methods are expected to produce similar results [122] and have a similar robustness [111]. In this thesis the method described in Ganapathy [44] is

used, along with a modification which was found to make the method more robust. There are also a number of calibration schemes which calculate $\mathbf{W}$ directly from a number of corresponding 2D and 3D point positions [98, 102, 122, 135]. Due to the non-linear nature of the rotational matrix, $\mathbf{R}$, an analytic solution has not been found for this problem and so all these papers use an iterative method.

Ganapathy [44] describes a closed form solution to the decomposition of $\mathbf{M}$ by forming a set of intermediate parameters which can be solved using linear equations. However, this method has been criticised by Weng *et al.* [135] and Phong *et al.* [98] because the orthogonality constraints associated with the rotation matrix $\mathbf{R}$ are not taken into account by the set of intermediate parameters. This results in the method being susceptible to noise [98]. I found the Ganapathy method [44] sensitive to errors in the matrix $\mathbf{M}$ which resulted in large errors in the decomposed rotation and translational parameters. Because of this, the method described in Ganapathy [44] was only used to calculate the intrinsic parameters. The extrinsic parameters were found using the following method. The *pseudo-inverse*, $\mathbf{P}^+$, of matrix $\mathbf{P}$ was calculated, see equation (4.8). This matrix satisfies the four *Moore-Penrose conditions* [47, 92]. If we multiply $\mathbf{M}$ by $\mathbf{P}^+$ we obtain a matrix which is very similar to the 4×4 rigid body transformation matrix $\mathbf{TR}$, as shown by equation 4.9. The difference between the two matrices, $\mathbf{P}^+\mathbf{PTR}$ and $\mathbf{TR}$ is that the fourth row of $\mathbf{P}^+\mathbf{PTR}$ is $(0\ 0\ 0\ 0)$, while the fourth row of $\mathbf{TR}$ is $(0\ 0\ 0\ 1)$. The fourth row of a 4×4 rigid body transformation matrix does not contain any information on the extrinsic parameters. Therefore, it is possible to decompose $\mathbf{P}^+\mathbf{PTR}$ to obtain the extrinsic parameters as if it was a standard rigid body transformation matrix. A rigid body transformation matrix is relatively easy to decompose, for example, from equation (4.6) element $r_{31} = \sin\theta_y$, therefore $\theta_y = \arcsin r_{31}$. I have found this approach to be an accurate and robust analytic method for calculating the intrinsic and extrinsic parameters from matrices created using the methods described in section 4.2.1.

$$\mathbf{P}^+ = \begin{pmatrix} 0 & 0 & 1 \\ k_1 & 0 & -c_s k_1 \\ 0 & k_2 & -l_s k_2 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.8}$$

$$\mathbf{P}^+\mathbf{M} = \mathbf{P}^+\mathbf{PTR} \tag{4.9}$$

4.3 Producing digitally reconstructed radiographs

This section discusses how to process the CT volume to produce a DRR. Sherouse *et al.* [108], Lemieux *et al.* [73] and Brown *et al.* [49] give descriptions on how to produce DRRs. The aim of this section is to describe in detail my method for producing DRRs and explain how my method differs from previous methods.

4.3.1 Interpolation in the CT volume

The matrix $\mathbf{M}$ is used to calculate the equations of 3D lines from the centre of each fluoroscopy pixel back to the x-ray source. The CT Hounsfield numbers along this ray path are integrated and the pixel value of the DRR set equal to this integral. However, the path of a given ray through the CT scan will traverse through a number of voxels by varying amounts. To calculate a value for the image intensity at a position which does not lie in the centre of a voxel requires image interpolation and sampling. There are a number of different interpolation techniques. These include, listed in terms of increasing computational cost, nearest neighbour, trilinear, tricubic and windowed sinc interpolation [54, 84, 96, 108]. As the interpolation method becomes more complex, typically more surrounding voxels are taken into account, nearest neighbour interpolation uses only a single voxel, while trilinear interpolation uses 8 voxels and tricubic uses 64 voxels. The number of voxels used by windowed sinc interpolation depends on the size of the windowing function, however they tend to be much more computationally expensive than tricubic interpolation [84]. The type of interpolation technique can have a large effect on the quality of the DRR and also on the time taken to produce a DRR. Because interpolation is carried out a large number of times along each ray path and there are a large number of rays (one for each pixel in the DRR), care has to be taken that the benefits from using a more complicated interpolation scheme (less blurring and aliasing) are not outweighed by the decrease in algorithm speed.

I have chosen to use bilinear interpolation which, with careful choice of sampling interval, approximates to trilinear interpolation. The method is as follows, where the 2D example, shown in figure 4.4, is used to show how bilinear interpolation can be approximated using linear interpolation. Figure 4.4 represents a 10x10 array of pixels where the centre of each pixel is positioned at the intercepts of the grid. The points at which the ray (bold line) enters and exits the pixel grid are calculated, positions A and B. The axis of the CT scan closest to the ray direction is chosen as the sampling axis. Therefore, in figure 4.4 the vertical axis is chosen as the sampling axis. The sampling interval is set to the pixel dimensions along the sampling axis projected onto the ray, e.g.

length d. Therefore, the sampling interval varies with respect to the angle between the ray and the sampling axis. The sampling positions along the ray are indicated using asterisks, *. Linear interpolation (bilinear in the 3D case) is carried out between the pixels either side of the ray, shown by dots, •, in the direction perpendicular to the sampling axis. When the integral along the entire ray path has been calculated it is then multiplied by d to account for an oblique ray path.

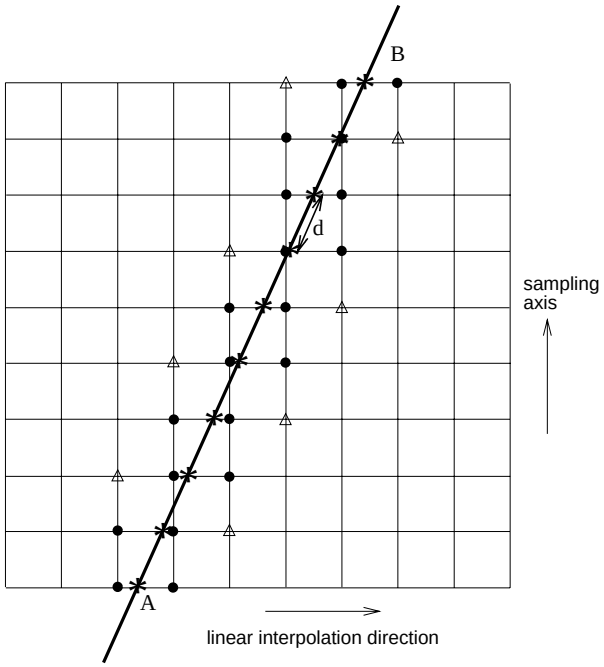
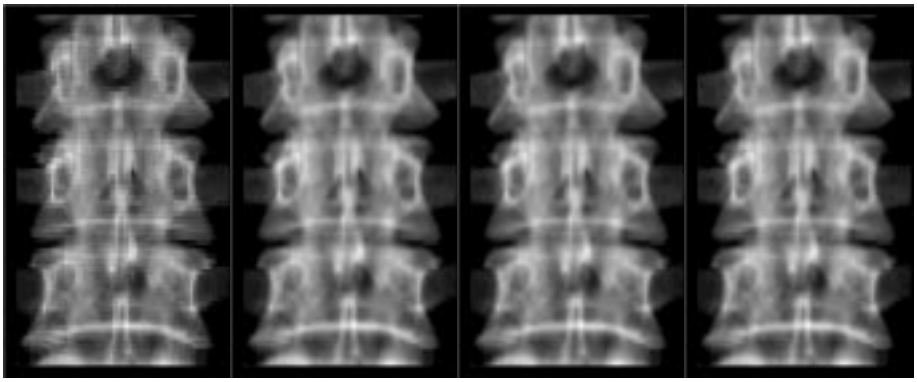


Figure 4.4: How bilinear interpolation can be approximated using linear interpolation. The grid represents a 10x10 array of pixels where the centre of each pixel is positioned at the intercepts of the grid. The figure shows exit and entry points A and B, sampling interval d, sampling positions *, voxels used in linear interpolation • and voxels which might have also been included if bilinear interpolation had been used △.

There is a small difference between the number of pixels used in the above mentioned method and the pixels used in standard bilinear interpolation. Standard bilinear interpolation could result (if different sampling positions were used) in the pixels marked by triangles (△) also being included in the interpolation process. These extra pixels become involved when the ray crosses over from one column of pixels to the next. However, these △ pixels are further away from the ray than the other pixels used in the interpolation process and so have the smallest influence on the interpolated value. This is shown in figure 4.5 where four DRRs have been produced from a CT volume of a spine phantom using different interpolation techniques; (a) nearest neighbour, (b) bilinear, and (c) and (d) use trilinear interpolation with different sampling intervals (d and d/2 respectively, see figure 4.4). It is easy to distinguish between the DRR produced using nearest neighbour inter-

polation (a) and the other three DRRs. However, the DRRs produced using bilinear and trilinear interpolation are visually indistinguishable. In fact, because the sampling rate used to produce images (b) and (c) was identical these images are identical.

In this thesis I have not compared the effects of different interpolation techniques on the performance of the registration algorithm. Instead I have opted to use bilinear interpolation as I believe it gives me the best trade off in the following two requirements: firstly, to keep the resampling artifacts to a minimum and secondly, to produce DRRs sufficiently quickly so that registration can occur in a clinically useful time.



(a) (b) (c) (d)

Figure 4.5: DRRs produced from a CT volume of a spine phantom using different interpolation methods, (a) nearest neighbour, (b) bilinear, (c) trilinear using the same sampling positions as used for the bilinear interpolation method and (d) trilinear using twice the sampling rate as used to produce (b) and (c). Note image (b) is exactly the same as image (c).

4.3.2 Avoiding truncation

Truncation occurs when rays cut through the top or bottom of the CT scan, see figure 4.6. Such rays are incomplete and should not be compared to the fluoroscopy image. Truncated rays were removed in the algorithm described in Lemieux *et al.* [73], while the other intensity based registration algorithms [45, 49] do not discuss the effects of truncation. In my algorithm truncated rays are detected and removed from the DRR.

Figure 4.7 shows a fluoroscopy image and a registered DRR. The effect of removing truncated rays can be seen by the straight edges at the top and bottom of the DRR shown in figure 4.7(b). The vertebrae in figure 4.7 are oriented so that their superior-inferior axis lies approximately parallel to the fluoroscopy screen. If the superior-inferior axis of the vertebrae was rotated either towards or away from the fluoroscopy screen (a θ_y rotation in figure 3.1) then the number of truncated rays would increase and so the amount of information in the DRR (e.g. the non-black pixels) would

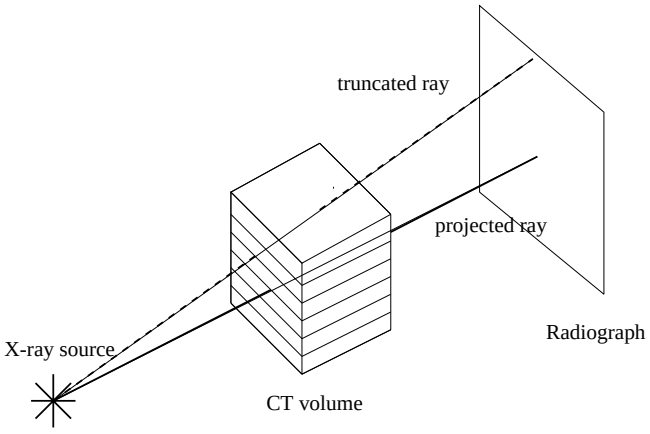
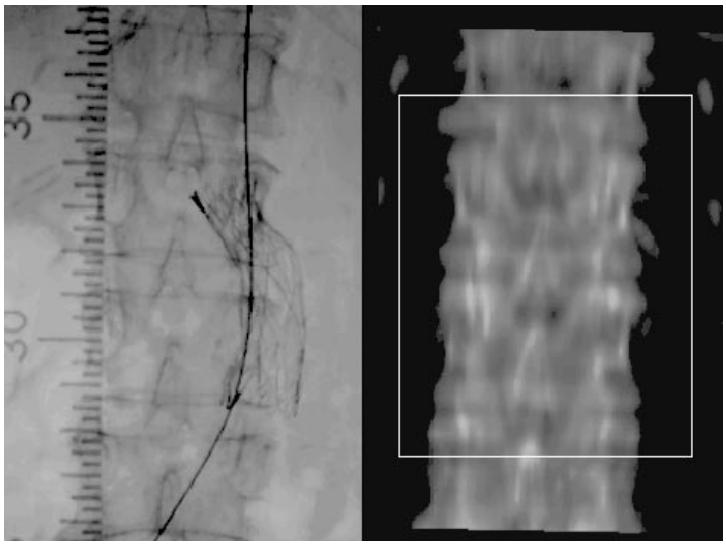


Figure 4.6: Diagram showing a truncated ray which should not be included in a DRR and a ray which should be projected and included in a DRR.

decrease. In some circumstances these alterations in the amount of information in the DRR can lead to misregistrations. One method to counter this problem is to define a region of interest (ROI) within the 2D image where truncation is unlikely to occur. An example of this is shown in fig. 4.7(b) where the white rectangle shows a ROI in the DRR image.



(a) (b)

Figure 4.7: Fluoroscopy image (a) and digitally reconstructed radiograph (b) at registration.

For the experiments described in this thesis regions of interest were manually defined taking the following considerations into account.

- The ROI should contain the feature of interest (usually a single vertebra) plus a small border region.
- The ROI should exclude areas where truncation was most likely to occur i.e. regions in the DRR which are produced from rays which intercept the CT volume close to its top or its base.

Using a ROI can have a number of additional benefits. Firstly, instead of casting rays through the CT volume to every pixel in the DRR, they are only projected to the ROI. This speeds up the production of DRRs. Secondly, because the registration is only carried out within this ROI, rather than the whole image, it should be more accurate within this region. Thirdly, if the ROI is defined so that it only contains a single rigid bony structure, then soft tissue deformation is not an issue.

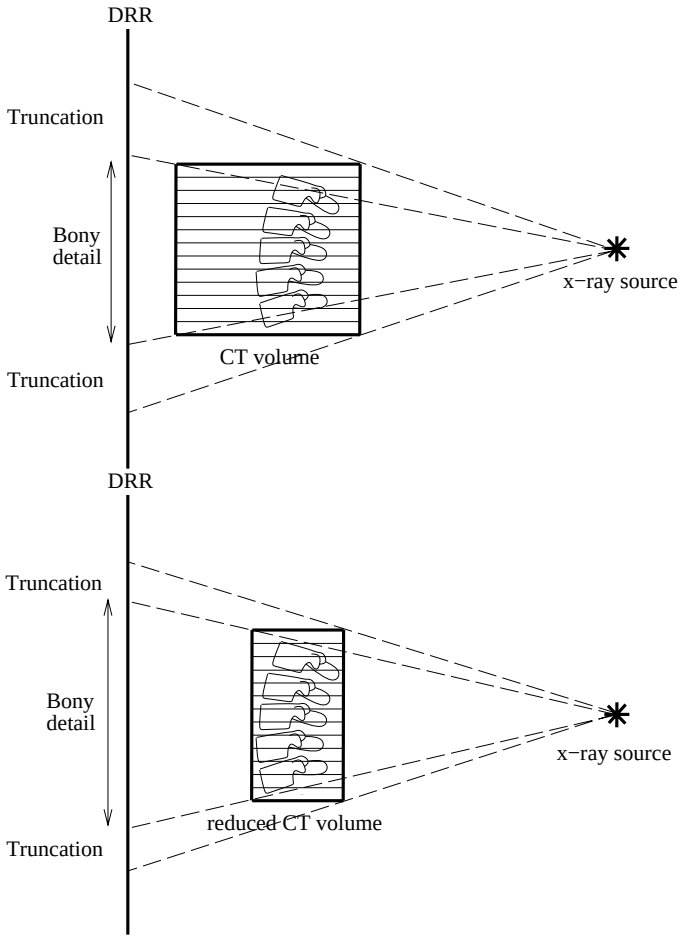


Figure 4.8: By reducing the in-plane dimensions of the CT volume the amount of truncation is reduced and the region of the DRR containing bony detail can be increased.

The amount of truncation depends on the ratio of in-plane image dimensions with respect to the height of the CT volume. If this ratio is large, then a large amount of truncation will occur. To decrease the amount of truncation a reduced CT volume is segmented from the entire CT volume. This reduced CT volume consists of a box surrounding the vertebrae, see figure 4.8. By reducing the in-plane dimensions the amount of truncation has been reduced and also the region of the DRR showing bony detail has increased. Because only voxels containing bone are projected (see section 4.3.3) no information is lost by reducing the size of the CT volume. There are a number of other benefits in using a reduced CT volume. Firstly, ray paths through the CT volume are shorter and so the time taken to produce a DRR is reduced. Secondly, the reduced CT volume can be defined so that the main feature of interest (usually a single vertebra) is placed approximately in the centre of the reduced CT volume. This has been found to be a useful method to further decouple the in-plane and out-of-plane translations. This is described in sections 5.4.3 and 6.5.

For the experiments described in this thesis reduced CT volumes were segmented from the entire CT volume. This segmentation process was achieved by manually defining a bounding box surrounding the vertebrae. The following points were taken into account when defining a bounding box.

- The feature of interest (usually a single vertebra) should be approximately in the centre.
- The in-slice dimensions of the box should be as small as possible, however, no voxels containing bone in the feature of interest should be removed.
- The size of the box perpendicular to the slice plane should be as large as possible to reduce the probability of truncation occurring.

4.3.3 Fluoroscopy pixel values and Hounsfield units

This section discusses the relationship between CT Hounsfield units and pixel intensities in fluoroscopy images. It is important to know this relationship so that DRRs can be produced which closely resemble fluoroscopy images.

An ideal relationship between pixel values in the DRR, I_{DRR}, and pixel values in the fluoroscopy image, I_{fl}, would be for them to be identical when the two images are registered; i.e. $I_{DRR}(i,j) = I_{fl}(i,j)$ over all the pixels i, j in the fluoroscopy and DRR images. However, such a direct relationship between pixel intensities would require detailed quantitative knowledge of parameters such as the effective energy of the x-rays used to produce the CT scan and the fluoroscopy image. These parameters are difficult to acquire on a routine basis as they often change

due to the imaging protocol. Because of these difficulties the algorithm produces DRRs so that, at registration, there should be a linear relationship between the DRR and the fluoroscopy pixel intensities, i.e. $I_{DRR}(i,j) = K_1(I_{fl}(i,j) + K_2)$ where K_1 and K_2 are constants. The rest of this section discusses how to achieve a linear relationship between fluoroscopy and DRR pixel intensities.

The value of a given fluoroscopy pixel depends on two main factors, the intensity of the incident x-rays on the image intensifier and the subsequent electronic processing which converts the x-ray intensity into a pixel value. Both of these factors are discussed in turn.

The intensity of an x-ray beam incident on a fluoroscopy screen depends on three main factors; the intensity of the x-ray source, the distance between the source and the image intensifier, and the attenuation of the beam due to interactions with material through which it passes. Assuming a monoenergetic beam of x-rays, the Lambert-Beer law for attenuation of radiation through a medium, equation (4.10), can be used to calculate the intensity of a beam of x-rays (I) which initially had an intensity, I_o, and has travelled a distance $\int dx$ through a medium which has a linear attenuation coefficient given by $\mu(x)$. The relationship between the Hounsfield value H, of a voxel and the linear attenuation coefficient of that voxel μ, is given in equation (4.1). These two equations can be used to calculate equation (4.11), which relates the intensity of x-rays incident on the image intensifier $I_{intensifier}$, to the integral of Hounsfield values along the path of the ray.

$$I = I_o e^{-\int \mu(x)\, dx} \tag{4.10}$$

$$I_{intensifier} = I_o e^{-\int \left(\frac{H(x)\mu_w}{1000} + \mu_w\right)\, dx} \tag{4.11}$$

The second factor concerns how the intensity of the incident x-rays on the image intensifier $I_{intensifier}$ is related to the pixel intensity in the fluoroscopy image I_{fl}. There are numerous electronic processing steps which take place between when an x-ray beam is incident upon an image intensifier and when the final output voltage or pixel value is displayed on the monitor. There can also be additional filters present, such as noise reduction and edge enhancement, which are used to improve the diagnostic image quality. I have not attempted to model these processes, but instead have assumed a logarithmic relationship between $I_{intensifier}$ and I_{fl}, see equation (4.12). This relationship is favoured in the production of digital subtraction angiography (DSA) images because it removes the modulation of image contrast by anatomy not containing contrast medium and so produces a uniform signal over the region containing contrast medium [67]. A logarithmic relationship also allows quantitative measurements to be taken, because the image

contrast is directly proportional to the projected width of the blood vessel. I believe that it is sensible to use a logarithmic relationship because digitally fluoroscopy systems are commonly specialised for the performance of DSA [4]. This relationship is also seen in the central portion of the characteristic curve in the production of radiographs [33].

$$I_{fl} \propto \log[I_{intensifier}] \tag{4.12}$$

By combining equations (4.11) and (4.12) it can be shown that I_{fl} is linearly proportional to the integral of $H(x)$ along the path of the ray, equation (4.13). Therefore, by integrating the Hounsfield numbers along the path of the ray the resultant pixel intensities in a registered DRR should be linearly proportional to the pixel intensities in a fluoroscopy image.

$$I_{fl} \propto \log[I_o] - \mu_w - \frac{\mu_w}{1000} \int H(x)\, dx \tag{4.13}$$

Equation (4.13) shows that there is a negative linear relationship between I_{fl} and $\int H(x)$. This is because an increase in the amount of attenuation along a ray path will result in a decrease in the intensity of x-rays incident on the image intensifier, which will reduce the pixel intensity in the fluoroscopy image. This is the opposite to a radiograph, where a decrease in the intensity of x-rays incident on the film, will result in the film remaining transparent and so, when viewed on a light box, the radiograph will appear bright.

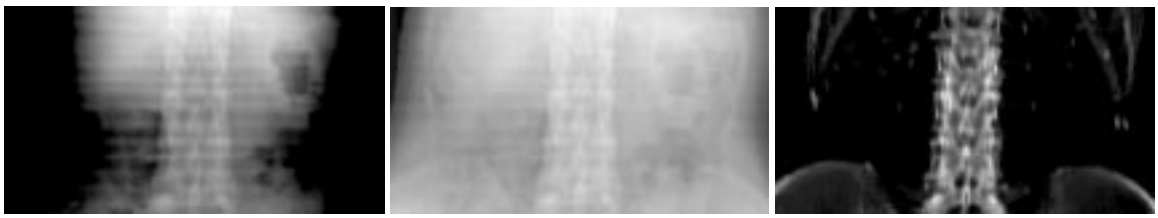
The above calculations have assumed that the fluoroscopy images and CT images are produced using x-rays which have the same effective energy. Generally this is not the case. The effective energy used to produce CT images is approximately 80keV, while the effective energy used to produce fluoroscopy images is between 50 to 60keV. The variation of attenuation coefficient with respect to x-ray energy is dependent on the material being irradiated. The ratio of the linear attenuation coefficient of bone to water at 50keV is larger than at 80keV, see equation (4.14). This results in a higher contrast between bone and soft tissue at 50keV compared to at 80keV. The effect on the DRR is that soft tissue structures swamp the signal from bone and no fine bony detail is visible. It is possible to account for the difference in effective energies (if they are accurately known) using the mixture rule which is outlined in appendix A.

$$\frac{\mu_{50keV}(bone)}{\mu_{50keV}(water)} > \frac{\mu_{80keV}(bone)}{\mu_{80keV}(water)} \tag{4.14}$$

I have opted for another solution, which is to threshold the CT volume so that no soft tissue structures are projected. The resultant DRR will contain only bony detail and so the reduction in

contrast between soft tissue and bone at the effective energy of CT x-rays is no longer an issue. At first sight it appears that by only projecting bony detail a large amount of useful soft tissue information is lost from the DRRs. However, I believe that for the applications in this thesis, registrations should not be carried out on the soft tissue structures because these structures can deform due to different patient positioning and due to the breathing motion of the patient. This violates the rigid body assumption and so can introduce errors into the registration.

Figure 4.9 shows the results of producing DRRs using three different methods; a straightforward integral of voxel intensities (a), using the mixture rule before integrating (b) and only integrating voxel intensities above a certain threshold (c). Note, the image produced using a straightforward integral, figure 4.9(a), required a large amount of windowing to show any bony detail, and features such as the ribs and pelvis are not visible in this image. No windowing was carried out on the images produced using the mixture rule and the threshold method. Both of these images show bony detail, such as the vertebrae, pelvis and ribs, though the threshold image shows much more detail than the image produced using the mixture rule.



(a) (b) (c)

Figure 4.9: DRRs produced using different methods to account for the differences in x-ray energies between CT images and radiographic images, (a) straightforward integral, (b) mixture rule, (c) only projecting voxels above 350 Hounsfield units. Note, image (a) has been windowed to improve contrast between bone and soft tissue, without this windowing no bony features would be visible. No windowing has been used in (b) or (c). Ribs and Pelvis are not visible in (a) while these features are just visible in (b) and clearly visible in (c).

4.4 Altering the fluoroscopy image

The previous section has outlined how to transform a CT volume into a DRR which is suitable for registering to bony features in a fluoroscopy image. For some differences between the modalities it is either necessary, or less computationally expensive, to alter the fluoroscopy image so that it more accurately depicts a DRR. These differences and the image processing techniques applied are outlined in the following sections.

4.4.1 Different resolutions between the modalities

Fluoroscopy images and CT images typically have different resolutions. Fluoroscopy images have square pixels which, for a 512×512 image produced using a 38cm image intensifier, gives a pixel size of 0.742×0.742mm. CT voxels typically have square in-slice dimensions, for example 1×1mm, but the slice thickness tends to be much larger; diagnostic scans of the abdomen typically range from a 3mm to a 10mm slice thickness. If an axial CT volume of the abdomen is compared to an AP fluoroscopy image then the direction normal to the slice plane will usually be approximately aligned with the vertical axis of the fluoroscopy image. The DRR shown in figure 4.7 was produced from a CT volume with a 3mm slice thickness. In this figure the DRR appears blurred when compared to the fluoroscopy image, especially normal to the CT slice plane. Bilinear interpolation was used to produce the DRR in figure 4.7, which helps to remove some of the effects of different resolutions.

Another method to reduce the effect of different resolutions is to blur the fluoroscopy image so its resolution resembles that of the DRR. Weese *et al.* [131, 134] and Brown and Boult [49] both blur the 2D image to make its resolution consistent with the CT voxel dimensions. Weese *et al.* [131, 134] use a 5×5 or a 3×3 neighbourhood averaging method. Brown and Boult [49] calculate the resolution limits of their DRR by assuming that their CT voxels are ellipsoids with principal axis lengths defined by the voxel dimensions. These ellipsoids were then projected, using orthogonal projection, into 2D to calculate the DRR resolution limits. The blurring method used by Weese *et al.* [131, 134] has the advantage that no assumptions are made about the orientation of the CT volume, and so the blurring cannot bias the algorithm to any particular direction. However, by using a square mask Weese *et al.* [131, 134] blur the images more than should be strictly necessary.

I use a similar method to Brown and Boult [49] to keep the amount of blurring to a minimum. The problem of introducing a bias is not believed to be large because a good estimate of the direction of maximum blurring can be obtained from the image data. In most cases the direction normal to the CT slice plane lies approximately parallel to the vertical axis in the fluoroscopy image. I use an elliptical mask with major, a, and minor, b, axes as given by equation (4.15). Two example ellipses for routinely used CT voxel and fluoroscopy pixel dimensions are shown in figure 4.10.

$$a = \frac{CT_z}{y_{pix}}, \quad b = \frac{CT_{xy}}{x_{pix}}, \tag{4.15}$$

where the in-plane voxel dimensions are assumed to be equal and are represented by CT_{xy} and the

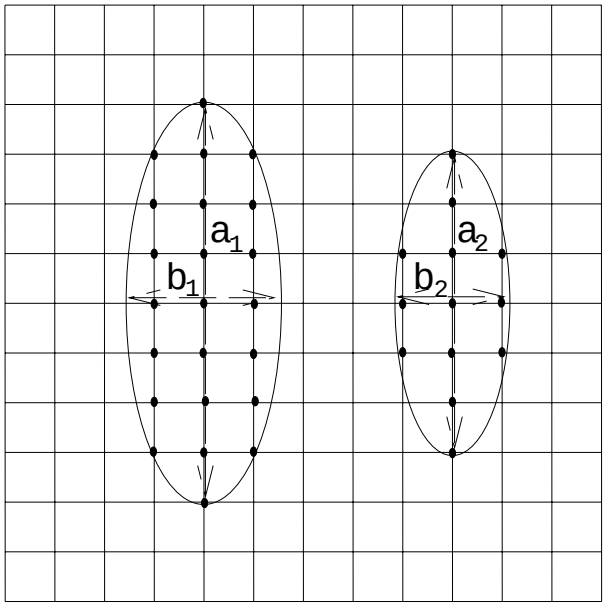


Figure 4.10: Two example ellipse sizes used for neighbourhood averaging. The grid represents an array of pixels where the pixel centres are at the intercepts of the grid. For both ellipses the example CT volume had in-plane voxel dimensions 1.094×1.094mm and a 3mm slice thickness. The larger ellipse represents the neighbourhood used for a fluoroscopy image containing 1024×1024 pixels and a 38cm field of view. The smaller ellipse represents the neighbourhood used for a fluoroscopy image containing 512×512 pixels and a 25cm field of view. In both cases the mask is passed over the image and the centre pixel is set equal to the mean value of all the pixels within the ellipse, represented by •'s

slice thickness is given by CT_z. The vertical and horizontal fluoroscopy pixel sizes are represented by y_{pix} and x_{pix} respectively. This elliptical mask is aligned so that the major axis is parallel to the vertical axis in the fluoroscopy image. The mask is passed over the fluoroscopy image and the value of the central pixel in the mask is set equal to the average value of all the pixels within the mask.

4.4.2 Distortion in the fluoroscopy image

There are two main types of distortion in fluoroscopy images, pincushion and "s-curve" distortion, these have been discussed in section 4.1.1. If only information from the centre of fluoroscopy images is used to carry out a registration then distortion may not be an issue. However, if information near to the edge of the fluoroscopy field of view is used, in particular, if fiducial markers are present near the edge of the fluoroscopy image, then pincushion and "s-curve" distortion can cause large registration errors.

Distortion correction has been carried out for all of the fluoroscopy images used in this thesis. A distortion correction phantom plus software, based on work by Haaker *et al.* [53], was made

available from Philips Medical Systems EasyVision Advanced Development. The phantom consists of an orthogonal array of steel spheres embedded in a rigid material with a low attenuation coefficient. The phantom is attached to the front of the image intensifier and an image acquired. The software automatically detects the centres of the spheres and uses a third order polynomial model to map the distorted grid into a regular orthogonal grid. The coefficients for the polynomial are found using a least-squares fit. Values of these coefficients were found for each of the magnification settings on the fluoroscopy sets for PA, lateral and oblique views. For subsequent images the most appropriate set of coefficients were chosen (based on magnification setting and image intensifier angle) and used to correct the image for distortion. These coefficients may vary with respect to time. In situations when accuracy was very important and access to the fluoroscopy machine was available the distortion correction phantom was imaged either prior to, or immediately after the subject being imaged.

4.5 Conclusions

In this section I have described in detail how the algorithm produces DRRs from the CT volume and also how the fluoroscopy image is processed so that it more accurately resembles a DRR. The novel contributions of my method are as follows. The algorithm thresholds the CT volume so that only voxels containing bone are projected. This results in the registration being based on bony detail and so soft tissue deformation is not an issue. I have used bilinear interpolation method, which produces approximately the same results as trilinear interpolation to produce good quality DRRs with a low computational cost. To reduce the number of truncated rays and to speed up the production of DRRs a reduced CT volume comprising of a box drawn around the vertebrae is used. Truncation effects can also cause misregistrations. To reduce the effects of truncation a ROI is defined within the 2D image and DRRs are only produced within this ROI. This results in a further speed up in the production of DRRs and also can be used to concentrate the algorithm on regions where the most accurate registration is required.

Chapter 5

A 2D-3D image registration algorithm

The previous chapter has described the information extraction step i.e. the formation of DRRs and the image processing applied to the fluoroscopy image to obtain two images which can be compared to each other. This chapter discusses the other three stages which combine to produce a registration algorithm, namely the similarity measure, search space and search strategy.

5.1 Changes in the imaged object

Many of the differences between the CT and fluoroscopy images have been been removed in the formation of DRRs and in the image processing applied to the fluoroscopy image. However, another set of differences between the images exists. These differences are not caused by the different modalities, but by changes in the imaged object. These differences will remain between the DRR and the fluoroscopy images, therefore, it is important that the similarity measure used by the algorithm is able to filter out these differences and so allow the registration to take place on information which is common to both images.

Overlying and underlying structures – As described in section 4.3.3 DRRs are created using a threshold so that no soft tissue structures are projected. However, soft tissue structures will still be present in the fluoroscopy image. For example, in figure 4.7 soft tissue structures can be seen in the fluoroscopy image but not in the DRR. Soft tissue structures are present in the majority of pixels and so can form a large difference between the images. Soft tissue structures tend to be larger in size and possess low spatial frequencies compared to bony structures. This characteristic can be exploited to design similarity measures which filter out the effect of soft tissue structures.

Interventional instruments in the field of view – During procedures interventional instruments are often introduced into the field of view of the fluoroscopy set. These instruments are usually made from materials with a high attenuation coefficient and so they can create large differences between the images. The difference is usually so large that these pixels contain very little useful information to aid the registration. Similarity measures should ideally, therefore, be insensitive to these highly altered pixels. Figure 4.7 shows three of these features: a ruler which is used to define positions during the procedure, a guidewire and a deployed aortic stent.

Spinal deformation – When using clinical images, the relative positions of the vertebrae may be different in the CT and fluoroscopy images. This may effect the registration. One method to reduce the effect of deformation is to register on a single rigid structure, such as a single vertebra [131]. Another method is to correct for the deformation by warping one of the images [75]. This will be discussed in more detail in chapter 8.

5.2 Similarity measures

The similarity measures in this thesis give an assessment of the registration between DRRs and fluoroscopy images. As mentioned in sections 4.3 and 4.4 the DRR has been created to closely resemble the fluoroscopy image and the fluoroscopy image has also undergone a number of image processing steps to increase its similarity with the DRR. However, a number of differences between the images will remain either because they are very difficult, or very computationally expensive, to correct for. There also exists a set of differences between the images which, no matter how accurately a DRR can be produced from a CT volume, would still exist. These are due to changes in the imaged object and have been discussed in the previous section.

This section describes seven similarity measures, six of which have been used previously for image registration and one new measure. In each case the measure is used to compare the pixel intensities in the fluoroscopy image, $I_{fl}(i,j)$, with the pixel intensities in the DRR, $I_{DRR}(i,j)$, where (i,j) defines the position of a pixel to be in column i and row j of the fluoroscopy image or DRR.

5.2.1 Normalised Cross Correlation

$$R = \frac{\sum_{(i,j)\in T}(I_{fl}(i,j) - \overline{I_{fl}})(I_{DRR}(i,j) - \overline{I_{DRR}})}{\sqrt{\sum_{(i,j)\in T}(I_{fl}(i,j) - \overline{I_{fl}})^2}\sqrt{\sum_{(i,j)\in T}(I_{DRR}(i,j) - \overline{I_{DRR}})^2}} \tag{5.1}$$

Where T is the region of overlap (i.e. where there are pixel values from both the DRR and the fluoroscopy image) and $\overline{I_{fl}}$ and $\overline{I_{DRR}}$ are the mean values of the images in the overlap region $(i,j) \in T$.

Normalised cross correlation is a standard statistical measure used to calculate whether two datasets are linearly related i.e. $I_{DRR}(i,j) = K_1(I_{fl}(i,j)+K_2)$ where K_1 and K_2 are a constants. The value of R lies between 1 and -1; a value of 1 is achieved if K_1 is positive and a value of -1 is achieved if K_1 is negative. As shown by equation (4.13), at registration, there should be negative linear relationship between DRR and fluoroscopy pixel intensities. Therefore, the algorithm will attempt to minimise R.

In normalised cross correlation the contribution from a given pair of pixels is strongly dependent on the intensity of the pixels. If there are differences between the images which create a small number of erroneous pixels that contribute significantly to the similarity measure then this can cause misregistrations. These can be caused by the presence of interventional instruments which greatly alter the image intensity of a small number of pixels in the fluoroscopy image. This measure has been used previously for 2D-3D image registration by Lemieux *et al.* [73].

5.2.2 Entropy of the difference image

The entropy measure H, described here operates on a single difference image I_{diff}, which is created by subtracting the DRR from the fluoroscopy image using a suitable scaling factor s, see equation (5.2).

$$I_{diff}(i,j) = I_{fl}(i,j) - s\, I_{DRR}(i,j) \tag{5.2}$$

$$H(s) = -\sum_x p(x) \log p(x) \tag{5.3}$$

The values from the difference image are separated into 64 evenly sized bins to form a histogram, and $p(x)$ in equation (5.3) denotes the probability of a pixel from I_{diff} belonging to bin x. Note that all the values in the difference image are used, not just the values in the overlap region. Entropy is maximised when all the probabilities $p(x)$ are equal and it is minimised if all the pixels in I_{diff} fall into a single bin.

The assumption when using entropy for 2D-3D image registration is that, as the registration position is approached, the difference image will contain less information and entropy will be minimised. At registration it should be possible to subtract the DRR from the fluoroscopy image so that corresponding features from the two images will vanish, as shown in figure 5.1.

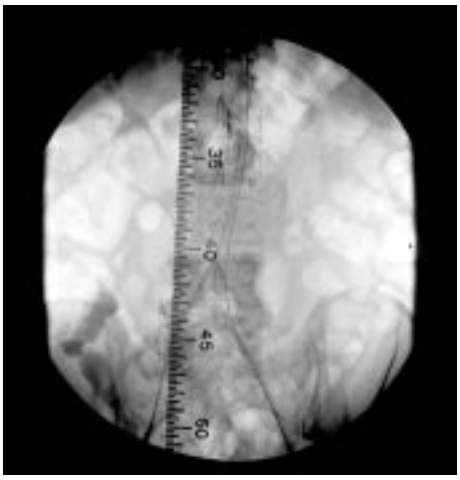


Figure 5.1: The result of subtracting a registered DRR of the L3 vertebra from the fluoroscopy image. Note that nearly all the structure from the L3 vertebra has been removed, while features such as the ruler and interventional guidewire are still clearly visible within the subtracted region.

To calculate a difference image the optimal scaling value s must be determined. The method used in this thesis begins with a scaling value of zero and then increases s in uniform steps until a minimum value of H is found. The size of the steps used in this thesis were equal to the ratio of the intensity ranges in the two images divided by 100.

An advantage of using entropy is that, because histograms are formed, the same weight is given to each pixel regardless of its intensity. This makes the measure much more robust to differences between the images which result in a small number of pixels having a large difference in intensity. This measure has been used previously in 2D-2D registration to correct for motion between images of a DSA sequence [18].

5.2.3 Mutual information

The calculation of mutual information S, like entropy, involves separating the pixel intensities into a number of bins. In mutual information three histograms are used, one for the fluoroscopy image, one for the DRR and a joint 2D histogram containing the pixel intensities from both images. The 1D histograms used contained 32 bins and the 2D histogram contains 32×32 bins. Initial experiments were carried out using 64 bins, however, the 64×64 2D histogram was found to be very sparsely filled which was thought to be the cause of some misregistrations. The probability distributions of the individual images are given by $p(x)$ and $p(y)$; and $p(x, y)$ represents the joint probability distribution. Unlike entropy, which works on a single difference image, mutual information only bins data in the overlap region between the images $(i, j) \in T$.

$$S = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{5.4}$$

Mutual information has been found to be very effective in 3D-3D multi-modality image registration between modalities such as MR and PET or MR and CT [81, 114, 115, 124]. Mutual information does not assume a linear relationship between the pixel values of the two images, but instead assumes that the co-occurrence of the most probable values in the two images is maximised at registration. This is a very useful property for 3D-3D multi-modality image registration. However, this property of mutual information may well be redundant for the purpose of registering a DRR to a fluoroscopy image because the DRR has been created so that, at registration, there should be a linear relationship between the pixel intensities in the DRR and fluoroscopy images.

5.2.4 Gradient correlation

Gradient based measures initially transform I_{fl} and I_{DRR} by differentiation. Horizontal and vertical 3x3 Sobel templates [48] were used to create gradient images $\partial I_{fl}/\partial i$, $\partial I_{fl}/\partial j$, $\partial I_{DRR}/\partial i$ and $\partial I_{DRR}/\partial j$ which represent the derivative of the fluoroscopy and the DRR images in the two orthogonal axes of the image, see figure 5.2. Normalised cross correlation, equation (5.1), is then calculated between $\partial I_{fl}/\partial i$ and $\partial I_{DRR}/\partial i$ and between $\partial I_{fl}/\partial j$ and $\partial I_{DRR}/\partial j$ within the region of overlap between the gradient images $(i,j) \in T_{\partial i}$ and $(i,j) \in T_{\partial j}$ respectively. The final value of gradient correlation, GC, is the average of these two normalised cross correlations.

$$GC = \frac{1}{2} \cdot \frac{\sum_{(i,j)\in T_{\partial i}} F_i D_i}{\sqrt{\sum_{(i,j)\in T_{\partial i}} F_i^2 \sqrt{\sum_{(i,j)\in T_{\partial i}} D_i^2}}} + \frac{1}{2} \cdot \frac{\sum_{(i,j)\in T_{\partial j}} F_j D_j}{\sqrt{\sum_{(i,j)\in T_{\partial j}} F_j^2 \sqrt{\sum_{(i,j)\in T_{\partial j}} D_j^2}}} \tag{5.5}$$

$$\text{where} \quad F_i = \frac{\partial I_{fl}(i,j)}{\partial i} - \overline{\frac{\partial I_{fl}}{\partial i}}, \qquad D_i = \frac{\partial I_{DRR}(i,j)}{\partial i} - \overline{\frac{\partial I_{DRR}}{\partial i}} \tag{5.6}$$

$$\text{and} \quad F_j = \frac{\partial I_{fl}(i,j)}{\partial j} - \overline{\frac{\partial I_{fl}}{\partial j}}, \qquad D_j = \frac{\partial I_{DRR}(i,j)}{\partial j} - \overline{\frac{\partial I_{DRR}}{\partial j}} \tag{5.7}$$

Gradient measures have the advantage that they filter out low spatial frequency differences between the images, such as those caused by soft tissue. They also concentrate the contributions to the similarity measure on edge information which intuitively appears sensible. Because this measure uses cross correlation it is expected to be affected by a few large differences in intensity. In particular the presence of interventional instruments in the fluoroscopy image, especially after they have been enhanced by the Sobel templates, may effect the performance of this measure.

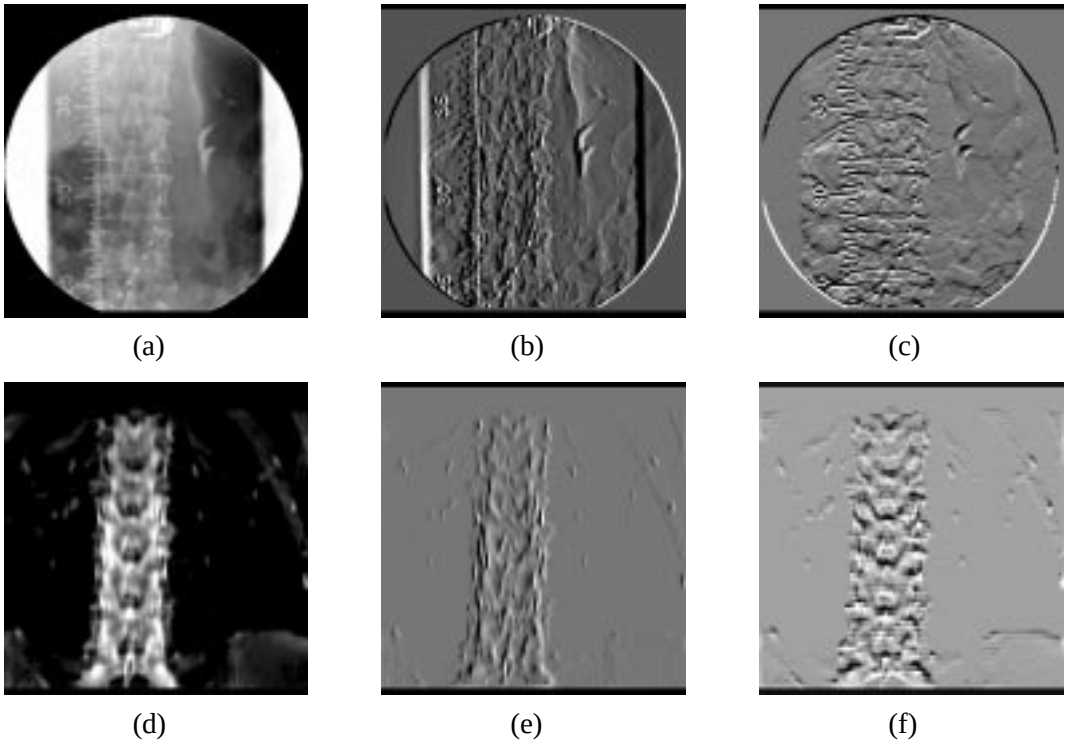


Figure 5.2: Gradient images formed using Sobel templates, (a) original fluoroscopy image, (b) dI_{fl}/di, (c) dI_{fl}/dj. (d) is a DRR which is registered to the fluoroscopy image, (e) dI_{DRR}/di and (f) dI_{DRR}/dj. Note that the intensity values in the fluoroscopy image have been inverted to make them consistent with the DRR.

A measure similar to gradient correlation has been previously used by Lemieux *et al.* [73] and Brown *et al.* [49].

Initial experiments with gradient measures were carried out using the modulus of the gradient i.e. $\|\nabla I_{fl}(i,j)\|$ and $\|\nabla I_{DRR}(i,j)\|$, see equations (5.8) and (5.9) where D_i, D_j, F_i and F_j are defined in equations (5.6) and (5.7). The measure based on the modulus of the gradient was far inferior in terms of accuracy and robustness compared to using the two individual gradients $\partial I_{fl}/\partial i$ and $\partial I_{fl}/\partial j$. By using the two gradients, the similarity measure is able to compare both the gradient intensity and the gradient direction.

$$\|\nabla I_{fl}(i,j)\| = \sqrt{\left(\frac{\partial I_{fl}(i,j)}{\partial i}\right)^2 + \left(\frac{\partial I_{fl}(i,j)}{\partial j}\right)^2} \tag{5.8}$$

$$\|\nabla I_{DRR}(i,j)\| = \sqrt{\left(\frac{\partial I_{DRR}(i,j)}{\partial i}\right)^2 + \left(\frac{\partial I_{DRR}(i,j)}{\partial j}\right)^2} \tag{5.9}$$

5.2.5 Pattern Intensity

The pattern intensity, P, similarity measure was developed by Weese *et al.* [131] at Philips Research Laboratory in Hamburg. While carrying out the work described in this thesis an amount of collaboration occurred between myself, who was sponsored by Philips Medical Systems, and Weese *et al.* [131]. However, I was not involved in the proposal or the initial development of this new similarity measure.

Pattern intensity operates on a difference image, I_{diff}, see equation (5.2). A circular mask, of radius r, is passed over the difference image and at each position the value of the central pixel in the mask (i, j) is compared to all the other pixels in the mask (w, v), using equations (5.10) and (5.11). This measure is a maximum when all the pixel values in the difference image are identical. At the registration position, when $s\, I_{DRR}$ is subtracted from I_{fl} to give I_{diff}, the structure from the vertebra should vanish. If there were no differences between the images then all the pixel values in the overlap region of the difference image would be identical and so the similarity measure would be maximised. However, there are differences between the two images, as shown by the soft tissue structures and interventional instruments visible in the overlap region in figure 5.1. To achieve accurate registrations despite the presence of these differences between the images the forms of equations (5.10) and (5.11) have been carefully chosen. Firstly, the regional nature of pattern intensity is able to reduce the effect of differences between the images which are on a larger scale than r. Therefore, differences with a low spatial frequency, such as soft tissue structures are filtered out, leaving the high spatial frequency bony detail on which the registration is based. Secondly the $1/(1 + x^2)$ form of equation (5.10) means that the maximum contribution to the similarity measure from any pair of pixels is the same. This helps to make the measure robust to differences between the images which introduce large differences in pixel intensity to a small number of pixels e.g. the presence of interventional instruments. The values for σ and r were chosen empirically in Weese *et al.* [131] to be $\sigma = 10$ and $r = 3$ pixels.

$$P_{r,\sigma}(s) = \sum_{i,j} \sum_{d^2 \le r^2} \frac{\sigma^2}{\sigma^2 + (I_{diff}(i,j) - I_{diff}(v,w))^2} \tag{5.10}$$

$$d^2 = (i - v)^2 + (j - w)^2 \tag{5.11}$$

5.2.6 Gradient difference

As mentioned in section 5.2.4 gradient based measures are potentially insensitive to soft tissue structures, but gradient correlation is expected to be sensitive to the presence of interventional

instruments. Because of this I propose another gradient based measure, gradient difference, see equation (5.12). It uses a difference image as explained in the entropy measure of section 5.2.2, though this time the difference image is calculated from gradient images, see equations (5.13). It employs the same $1/(1+x^2)$ form as pattern intensity which should make the measure robust to the presence of interventional instruments.

$$G(s_{_V}, s_{_H}) = \sum_{i,j} \frac{A_v}{A_v + (I_{diffV}(i,j))^2} + \sum_{i,j} \frac{A_h}{A_h + (I_{diffH}(i,j))^2} \tag{5.12}$$

$$\begin{aligned} I_{diffV}(i,j) &= \frac{\partial I_{fl}}{\partial i} - s_{_V} \frac{\partial I_{DRR}}{\partial i} \\ I_{diffH}(i,j) &= \frac{\partial I_{fl}}{\partial j} - s_{_H} \frac{\partial I_{DRR}}{\partial j} \end{aligned} \tag{5.13}$$

The gradients in equation (5.13) are described in the gradient correlation section 5.2.4, and A_v and A_h in equation (5.12) are constants, which for these experiments were the variance of the vertical and horizontal gradient fluoroscopy images.

5.2.7 Local correlation

This measure uses the same regional concept as pattern intensity, but uses normalised cross correlation to compare the pixel values within a given region. By using the regional method the measure should be more robust to soft tissue structures. Also the effect of interventional instruments should be confined to within the region r and so should not have a significant effect on the overall similarity measure. The use of normalised cross correlation, rather than equation (5.10), means that the measure does not need to optimise the scaling parameter s. However, because cross correlation can be greatly effected by a single pixel value, an interventional instrument may effect the performance of local correlation for all pixels within a radius r of the interventional instrument. This reduces the number of useful pixels available to the similarity measure. This measure has been used previously by Collins *et al.* [29] to carry out a non-linear mapping between MR brain images and a human brain model.

$$L = \sum_{(i,j)\in T} \frac{\sum_{d^2\leq r^2}(I_{fl}(u,w) - \overline{I_{fl}(i,j)})(I_{DRR}(u,w) - \overline{I_{DRR}(i,j)})}{\sqrt{\sum_{d^2\leq r^2}(I_{fl}(u,w) - \overline{I_{fl}(i,j)})^2}\sqrt{\sum_{d^2\leq r^2}(I_{DRR}(u,w) - \overline{I_{DRR}(i,j)})^2}} \tag{5.14}$$

$$d^2 = (i-v)^2 + (j-w)^2, \tag{5.15}$$

where $\overline{I_{fl}(i,j)}$ and $\overline{I_{DRR}(i,j)}$ represents the mean pixel value within a radius r of pixel (i,j) in the fluoroscopy and the DRR images respectively.

5.2.8 Discussion on similarity measures

The above similarity measures can be split up into a number of groups based on common attributes. For example, whether the measures are applied to the whole image or just the region of overlap, whether they are applied to pixel intensities or the gradient of intensities and what relationship is assumed between the features in each image.

In this thesis the measures have been split up in the following two ways; firstly based on what information from the images are compared by each measure. The measures are split up based on whether they compare pixel intensities or the gradient of pixel intensities. The measures are also split up based on whether all the image intensities are combined into one statistical measure, or only intensities within a local region are compared, see table 5.1.

The second way in which the measures are split up depends on whether the maximum contribution from a given pair of pixels to the similarity measure is dependent or independent of the pixel intensities (table 5.2). The sets of similarity measures have been labelled types 1, 2, 3, A and B so that they can be referred to in the following chapters.

Compared Features	Pixel intensities		Gradient of pixel intensities
Region of comparison	local region	Whole image or area of overlap	
Measures	local correlation pattern intensity	cross correlation entropy mutual information	gradient correlation gradient difference
Type	1	2	3

Table 5.1: Three types of similarity measure based on what information is compared by each measure and within what area of the image are they compared.

Maximum contribution from a pair of pixels	independent of intensity values	dependent on intensity values
Measures	Entropy Mutual information Pattern intensity Gradient difference Local correlation	Cross correlation Gradient correlation
Type	A	B

Table 5.2: Two types of similarity measure based on whether or not the maximum contribution to the similarity measure from a given pair of pixels is independent of the pixel intensities.

5.3 Search space

The search space chosen for the algorithm is the six rigid-body degrees of freedom. The intrinsic parameters are calculated prior to registration using a calibration object (see figure 3.2). A full ten degree of freedom registration was not attempted for three main reasons. Firstly there is a large body of literature available on how to accurately obtain intrinsic parameters [19, 44, 98, 102, 122, 135]. These methods are able to use precisely machined calibration objects and they can calculate the intrinsic parameters to a high degree of accuracy.

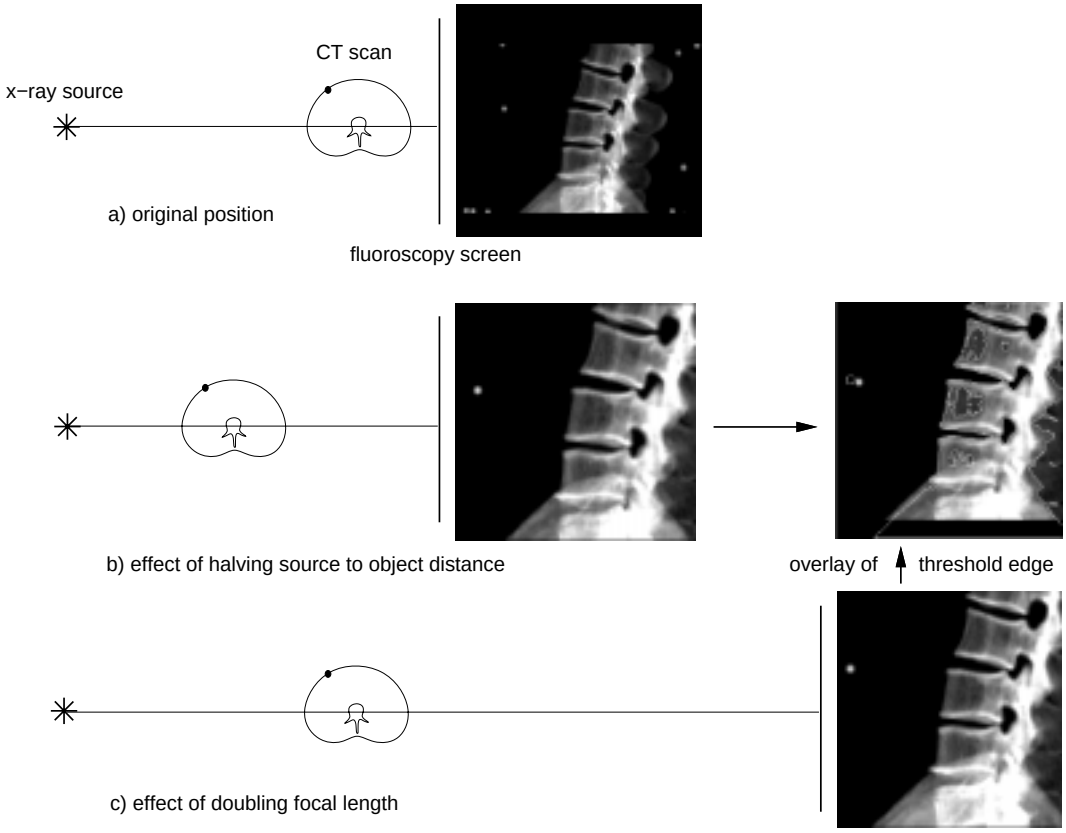


Figure 5.3: The difference between altering the source to object distance and increasing the focal length. (a) shows the original position for reference; (b) shows the result of halving the source to image distance; c) shows the result of doubling the focal distance; also shown is a threshold overlay of c) onto b).

Secondly, the alteration of some of the intrinsic parameters can have nearly exactly the same effect on the final DRR as altering some of the extrinsic parameters. A good example is to compare translations along the source-to-film axis with alterations in the focal length. This is illustrated in figure 5.3 where lateral DRRs have been produced of a lumbar spine phantom using three different sets of intrinsic and extrinsic parameters. The original image is shown in (a) and images (b) and (c)

show the effect of halving the source to object distance and doubling the focal length respectively. Although images (b) and (c) appear similar they are different. Image (c) is simply a magnification of image (a). While the change between image (a) and image (b) is more complex. This is because in producing image (b) the distance between the x-ray source and the object has altered and so the paths of the rays through the CT volume have also changed. Figure 5.3 also shows a threshold overlay of image (c) onto image (b). The outline of the vertebrae match almost exactly. This shows how small the difference is between these two DRRs, even though there is a very large difference in the intrinsic and extrinsic parameters used to produce these two images. This occurs because the outlines of the vertebrae lie approximately in a plane which is parallel to the imaging plane. It would be extremely difficult for a similarity measure to detect the difference between altering the focal length and altering the source to object distance. Calculating the correct parameters is important because, although the bony features on which the registration is based may appear registered, errors can occur if other information (which is in a different plane to the vertebrae) is projected. An example is shown in figure 5.3 by the projection of a fiducial marker, shown by the black dot on the outline of the CT slice and by the white dot in the DRRs. Differences in the position of this marker can clearly be seen in the overlay image. There is also a similar trade off between altering the in-plane translational parameters and altering the intrinsic parameters c_s and l_s.

The third reason for limiting the search space to six degrees of freedom is the computational cost involved in optimising a large number of parameters. In going from six to ten parameters the size of the search space greatly increases with an associated increase in registration time. The actual increase in registration time depends greatly on the individual registration scenario, the behaviour of the similarity measure within the search space and the search strategy used. For example, let us assume a situation where the starting estimate is within n steps of the registration position in each degree of freedom. As a worst case scenario, where a global search strategy is used, the number of calculations of the similarity measure would increase by a factor of n^4 if the search space was increased from 6 to 10 degrees of freedom. However, as a best case scenario, if a number of successive line minimisations could be carried out, then the number of calculations of the similarity measure might only increase by a factor of $\frac{10}{6}$. The algorithm described in this thesis uses a gradient descent search strategy which is described in the following sections. In a best case scenario using this search strategy, where the algorithm takes the minimum number of steps required to reach the registration position, I estimate that the number of calculations of the similarity measure would increase by a factor of between 2 and 3 if the search space was increased

from 6 to 10 degrees of freedom. However, the extra size and complexity of the search space may result in the algorithm taking many more than the minimum number of steps required to reach the registration position, which would in turn further increase the registration time.

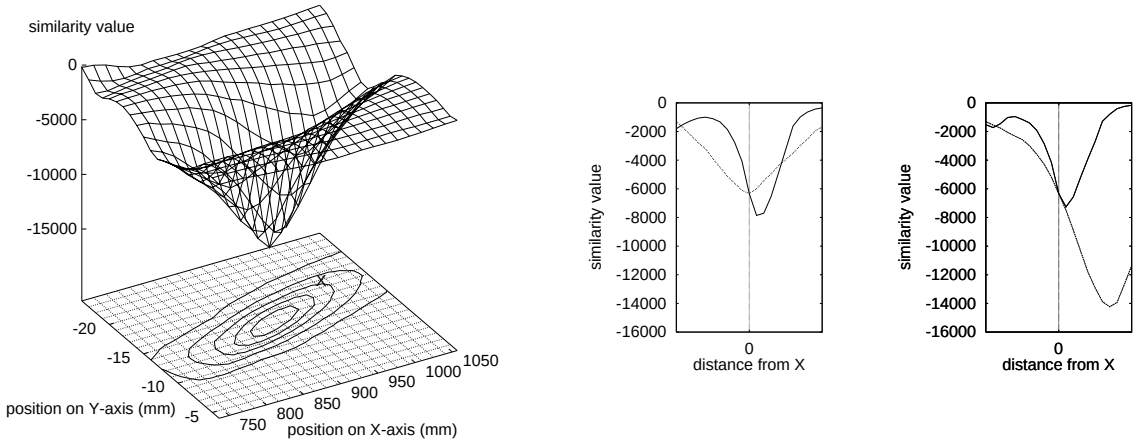
5.4 Search Strategy

Search strategy defines how the algorithm moves through the search space to reach an optimum value of the similarity measure. This section is split up into three parts. The first part discusses the problems involved with attempting to visualise a six dimensional search space using 1D or 2D plots of the similarity measure. This is followed by a discussion on how the extrinsic parameters are altered to reach registration. The final part of this section discusses the choice of axes about which translations and rotations occur.

5.4.1 Visualising the behaviour of the similarity measure

The search strategy chosen depends greatly on the behaviour of the similarity measure within the search space. To rigorously investigate the behaviour of the similarity measure an n dimensional plot should be produced where n is the number of degrees of freedom in the search space. Analysing and producing a six dimensional plot is extremely difficult and so commonly only one or two dimensional plots are used to give a visual impression of the behaviour of the similarity measure. This is fraught with difficulties and can have serious drawbacks as the following example illustrates. Figure 5.4(a) shows a plot of a two dimensional search space which consists of a long narrow valley with a minimum (the registration position) in the centre of the plot. If we consider ourselves currently at position X on this graph and try to represent the 2D search space by 1D plots through position X, then the accuracy of this representation depends greatly on the directions chosen for the 1D cuts. If the cuts are parallel to the X-axis and Y-axis shown in the 2D plot, then the resultant 1D plots are shown in figure 5.4(b), where the solid line represents a cut along the direction of the X-axis and the dashed line represents a cut along the direction of the Y-axis. From figure 5.4(b) position X appears to be close to registration while, as seen from the 2D plot, this is clearly not the case. Figure 5.4(c) shows two orthogonal 1D plots, one along an axis which is close to the major axis of the valley, shown by the dashed line, and the other at right angles to the major axis, shown by the solid line. This graph shows a much better representation of the 2D plot, where position X is close to the minimum position in one direction but a long way from the minimum position in the other direction. This example emphasises the problems which can arise

when a 2D search space is represented by two 1D graphs. Far greater problems can occur when representing 6D search spaces by 1D or 2D plots. Therefore, care should be taken when using 1D or 2D plots to help visualise the performance of a similarity measure in a 6D search space.

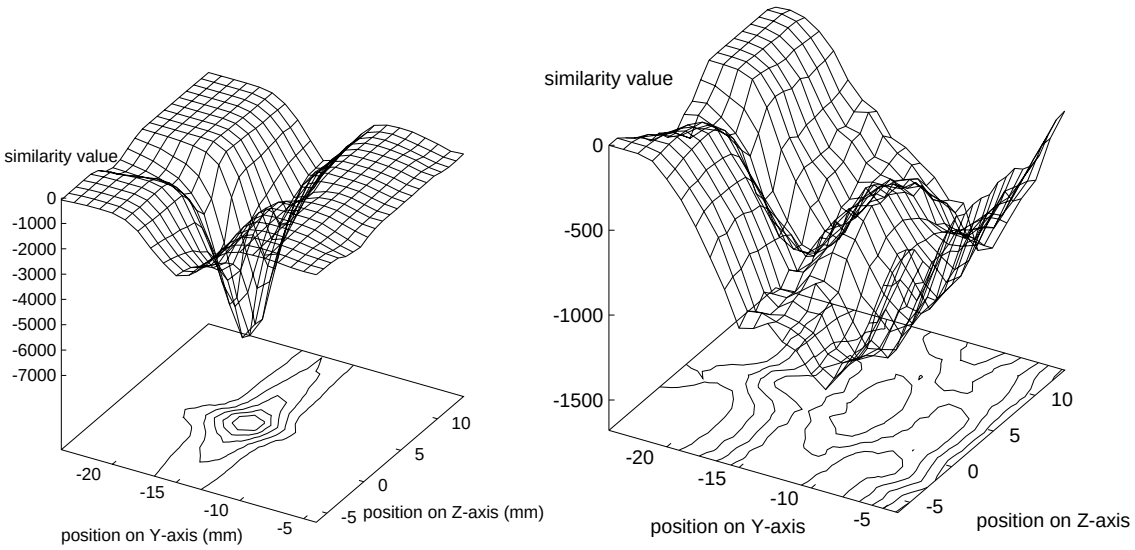


(a) (b) (c)

Figure 5.4: The effect of representing a 2D search space (a) with two one dimensional orthogonal cuts. In (b) the cuts were along the direction of the X-axis (solid line) and Y-axis (dashed line). In (c) the axes were aligned parallel to (dashed line) and at right angles to (solid line) the major axis of the valley. Note: from (b) position X appears to be very close to registration, while from (a) and (c) this is clearly not the case. This illustrates the dangers involved in using plots which are of a lower dimensionality than the search space.

5.4.2 The choice of search strategy

Bearing in mind the problems outlined in the previous section a number of 2D plots have been produced to help explain the search strategy chosen for the registration algorithm. The plots show how a similarity measure (gradient difference) behaves when comparing a fluoroscopy image to DRRs of a spine phantom which is described in detail in section 6.1. Figure 5.5(a) shows the effect on the similarity measure of varying the in-plane translational parameters Y and Z. The other four extrinsic degrees of freedom are set at the registration position. This plot shows a definite minimum of the similarity measure in the centre of the plot which is the known registration position. Figure 5.5(b) has been produced to show how the similarity measure behaves, again when Y and Z are altered, but this time the other four extrinsic parameters are not at the registration position, instead they have been moved from the registration position to a typical starting position for the algorithm. The plot shown in this figure is much rougher, there are a number of local minima and the minimum value is not at the registration position. If an optimisation scheme which used successive line minimisations, such as Powell's multidimension set method [100] was used



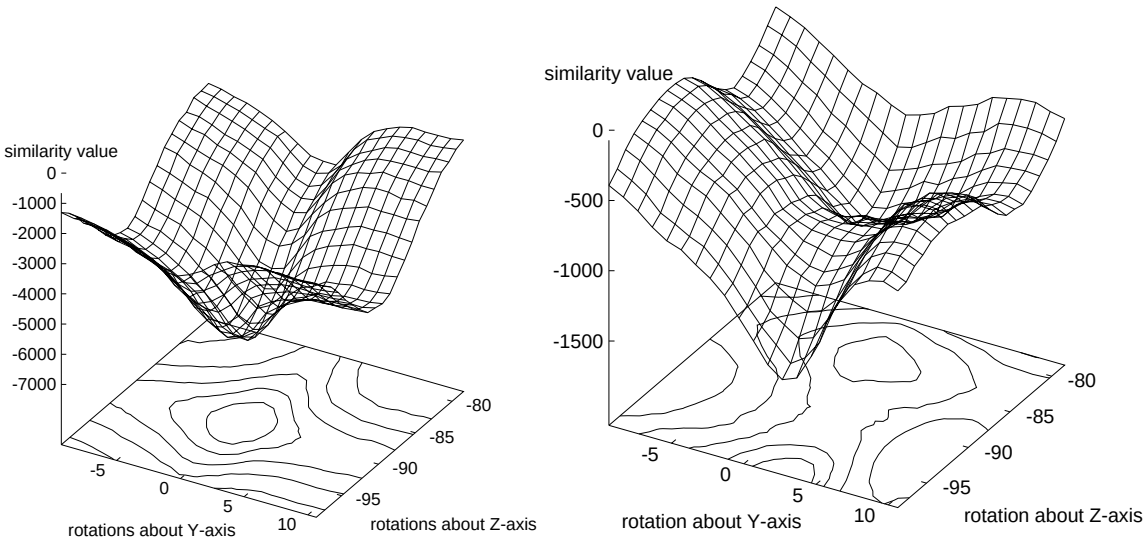
(a) (b)

Figure 5.5: The behaviour of a similarity measure while varying the two in-plane translational parameters where (a) the other four parameters are at the registration position and (b) the other four parameters are at a typical starting position for the registration algorithm

to try and find the minimum value of the similarity measure in figure 5.5(b), then the algorithm could easily wander a large distance from the registration position.

Figure 5.6 shows the effect of altering the out-of-plane rotational parameters. In figure 5.6(a) the other four extrinsic parameters are at the registration position. This plot shows a definite minimum at the known registration position. When the other four extrinsic parameters are changed to represent a typical starting position for the algorithm the plot is greatly altered, as shown in figure 5.6(b). The global minimum is now a long distance from the registration position, though a local minimum does still exist fairly close to registration position. Again if successive line minimisations were used to find this local minimum then, depending on the starting position, the algorithm could wander a large distance from registration.

These examples show that, at typical starting positions, values of a similarity measure do not plot out a smooth graph within the search space and there can be a large number of local minima. Because of this, if a search strategy which employs a number of 1D line minimisations is used, then the algorithm could wander a long distance from the registration position. A search strategy is required which either moves more methodically through the search space, such as a gradient descent method or the simplex algorithm [100], or a more global approach is required, such as simulated annealing or genetic algorithms. For this algorithm we have chosen the former type of strategy, firstly because it is more reproducible and secondly because it typically requires fewer



(a) (b)

Figure 5.6: The behaviour of a similarity measure while varying the two out-of-plane rotational parameters where (a) the other four parameters are at the registration position and (b) the other four parameters are at a typical starting position for the registration algorithm

iterations to reach registration.

The search strategy used by the algorithm is a gradient descent type method. The method is as follows.

- The current set of extrinsic parameters, $\mathbf{E}$, are used to produce a DRR.
- This DRR is compared to the fluoroscopy image to obtain the current value of the similarity measure, (V_0).
- The extrinsic parameters are altered, one at a time, by $\pm\omega_{\mathbf{E}}\delta\mathbf{E}$ and a new DRR is produced at each position. Note, $\delta\mathbf{E}$ is the current step size and $\omega_{\mathbf{E}}$ is a weighting factor.
- These DRRs are compared to the fluoroscopy image to obtain 12 values for the similarity measure, $(V_1, V_2, \ldots, V_{12})$.
- The values $V_0, V_1, \ldots, V_{12}$ are used to determine which extrinsic parameters currently have a local optimal value and which parameters could be changed to improve the value of the similarity measure.
- The parameters which improve the value of the similarity measure are altered. These parameters are altered by an amount, based on the size of the improvement in the similarity

measure, where the largest alterations occur for the parameters with cause the largest improvement in the value of the similarity measure.

- The above steps are repeated until all six parameters are at an optimum value ($V_0 < V_i, i = 1, \ldots, 12$). Then the step size is reduced and the registration process is continued until a minimum value for the step size is reached.

The weighting factor $\omega_{\mathbf{E}}$ equals unity for all the extrinsic parameters except for translations in the X direction. Because movements along X have a much smaller effect on the DRR than movements in the other degrees of freedom, $\omega_{_X} = 4$. This value was chosen empirically. The factors involved were to ensure that the step size was large enough to avoid noise in the value of the similarity measure when optimising the X parameter, but small enough so that the parameter could still be optimised relatively accurately.

The algorithm uses a multi-resolution approach with a number of different step sizes. The fluoroscopy image and the CT volume are reduced in size by blurring and subsampling. Two different sizes of fluoroscopy images are used, 256×256 and 128×128 pixel images. The CT volume is reduced to a size so that no subsampling occurs when the DRRs are produced. This is achieved by calculating the distance d_{ray}, between adjacent rays at the centre of the CT volume. The voxel sizes along each axis of the CT volume were taken in turn and compared to d_{ray}. If the voxel size was smaller than d_{ray}, then the voxel size in this direction was doubled and the resolution was halved by blurring and subsampling. This process was repeated until d_{ray} was less than the voxel size. By reducing the number of pixels in the fluoroscopy image and the number of voxels in the CT volume the time taken to produce a DRR was greatly reduced.

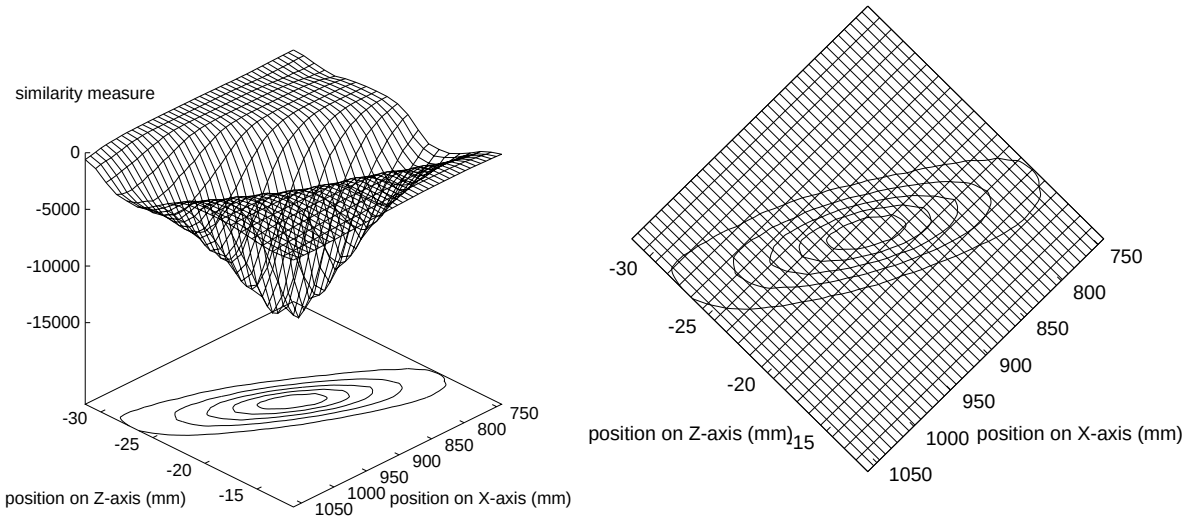
Table 5.3 shows the seven different combinations of step size and fluoroscopy image size used by the algorithm. The algorithm begins with 128×128 pixel images and $\delta\mathbf{E}$ is set equal to 4 mm or 4 degrees. When an optimum value is found in all six parameters $\delta\mathbf{E}$ is halved until $\delta\mathbf{E} = 0.5$mm or degrees. At this point a finer resolution is used (256×256 pixel images) and $\delta\mathbf{E}$ is set to 2mm or degrees. Again, when an optimum value is found in all six parameters $\delta\mathbf{E}$ is halved until $\delta\mathbf{E} = 0.5$mm or degrees. The multi-resolution approach was found to greatly reduce the time taken to reach registration and helped avoid local minima. This optimisation scheme was used as it is relatively simple and should be robust. A similar optimisation scheme has been found to be robust for 3D-3D registration [114, 115].

Step size (mm or degrees)	Fluoroscopy image size (pixels)
4	128×128
2	128×128
1	128×128
0.5	128×128
2	256×256
1	256×256
0.5	256×256

Table 5.3: The seven different combinations of step size and fluoroscopy image size used by the algorithm.

5.4.3 Rotation and translation axes

When attempting to optimise a similarity measure it can be useful to find a set of "non-interfering" or conjugate directions; this is the basis behind Powell's multidimension set method [100]. This is particularly important if the values of the similarity measure form a long narrow valley within the search space. Such a situation is commonplace in 2D-3D registration because alterations in the in-plane extrinsic parameters can have a much larger effect on the appearance of the DRR than alterations in the out-of-plane extrinsic parameters. As described in section 3.1.3, the most marked example is comparing an in-plane with an out-of-plane translation.



(a) (b)

Figure 5.7: The creation of local minima because in-plane and out-of-plane translations are not decoupled.

Figure 5.7 shows a plot of the gradient difference similarity measure when comparing a fluoroscopy image to DRRs produced from a CT scan of a spine phantom. The plot maps out

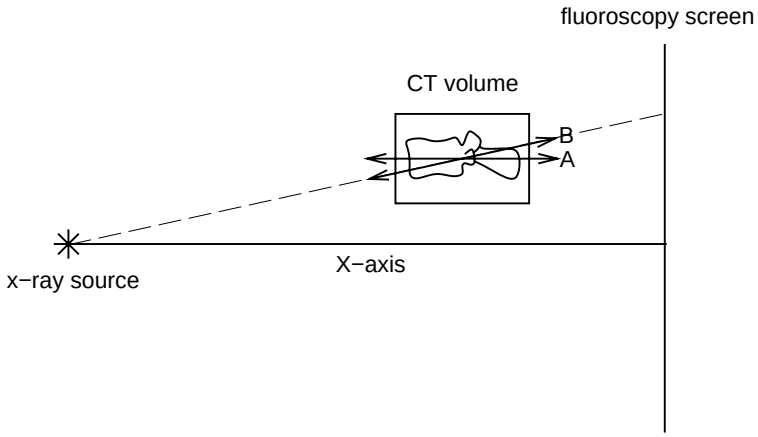
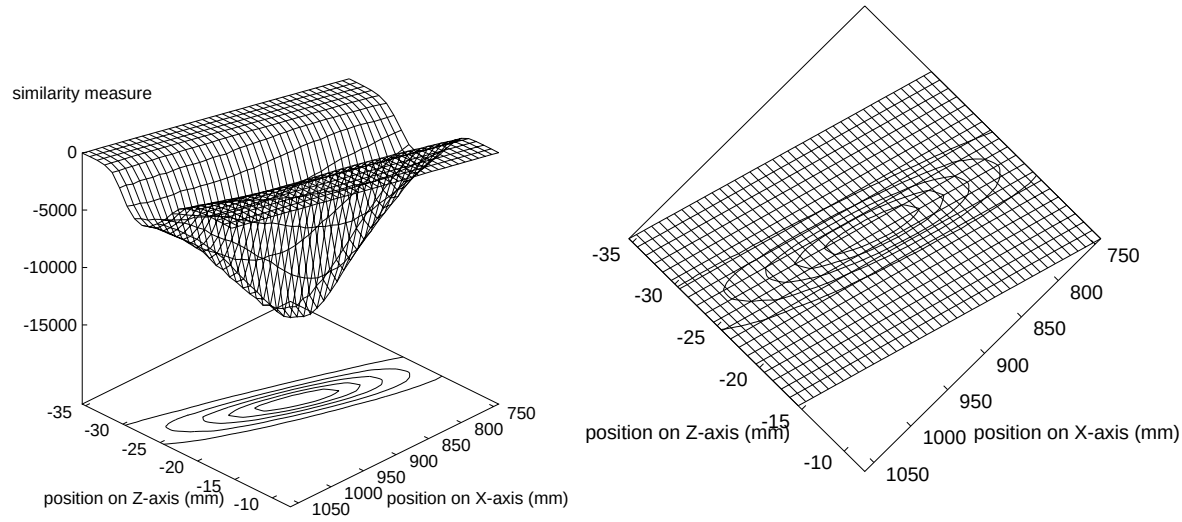


Figure 5.8: Showing the two different out-of-plane translation directions. Direction A is parallel to the X-axis, while direction B is along a line from the x-ray source to the centre of the CT volume. Direction B is preferable as it helps to decouple the in-plane and out-of-plane translational parameters.

the behaviour of the similarity measure as the out-of-plane translational parameter (translations parallel to the X-axis in figure 3.1) and an in-plane translational parameter (translations parallel to the Z-axis in figure 3.1) are altered. The plot shows that the registration position, the minimum point, is in the centre of a long narrow valley. The axes used to produce figure 5.7 do not form a set of conjugate directions because translations parallel to X-axis will, unless that axis goes through the centre of the CT volume, also result in an in-plane translation. If this resultant in-plane translation causes a misregistration which has a larger effect on the similarity measure than the out-of-plane translation moving towards registration, then a situation can arise where the out-of-plane parameter cannot be optimised. This effect is seen as the formation of local minima around the registration position in figure 5.7(a). Also, if figure 5.7(a) is viewed from above, as shown by figure 5.7(b), the axes can be seen to lie at approximately 45^o to the long axis of the valley. If these axes were used by the registration algorithm then, to reach the registration position the algorithm must zig-zag down the valley which would require a large number of iterations. Because of these problems, the out-of-plane translation direction used by the algorithm is not the X-axis, but instead it is along a line from the x-ray source to the centre of the CT volume, see figure 5.8.

Translations along a line from the x-ray source to the centre of the CT volume do not cause an in-plane translation. The result of this modification can be seen in figure 5.9. No local minima can be seen in figure 5.9(a) and now the directions in which the algorithm searches line up with



(a) (b)

Figure 5.9: How the in-plane and out-of-plane parameters can be decoupled by translating along an axis connecting the x-ray source to the centre of the reduced CT volume.

the major and minor axes of the valley, see figure 5.9(b).

There could also be advantages in altering the direction of the in-plane axes. For example, if the object to be registered is elongated, such as a femur or a humerus, then it may be sensible to define a set of axes based on the principle axes of the object, instead of using a set of axes defined by the fluoroscopy gantry.

5.5 Conclusions

This chapter has described three of the four major components which combine to produce the registration algorithm, namely similarity measure, search space and search strategy. I have described seven similarity measures one of which has not been previously used for 2D-3D registration (local correlation) and one new measure, gradient difference. The similarity measures have been placed into a number of groups which will be referred to in the following chapters. The search space for the algorithm is defined by the six extrinsic parameters and I have described in detail why the algorithm does not also include the intrinsic parameters in the search space. The search strategy for the registration algorithm is a gradient descent method. Also, the out-of-plane translation axis for the algorithm is not orthogonal to the in-plane translational axes, as is typically used by 2D-3D registration algorithms. Instead the out-of-plane translation axis is along a line from the x-ray source to the centre of the CT volume. This direction has been chosen because it decouples the in-plane and out-of-plane translational parameters which should help to avoid local minima.

Chapter 6

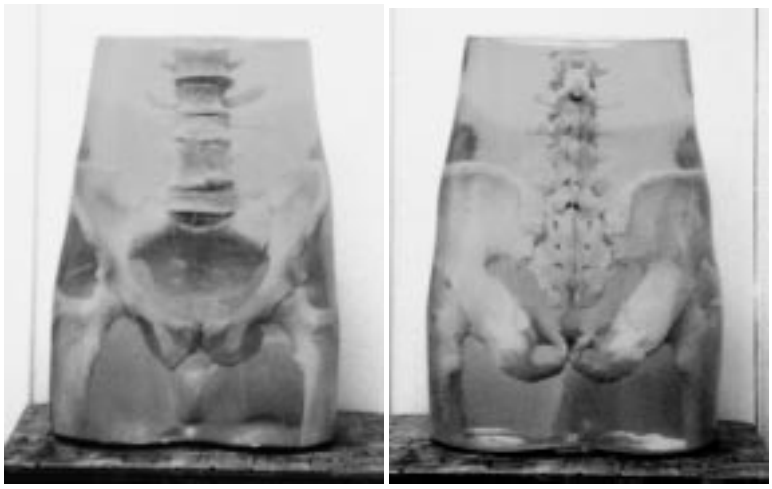
An experimental comparison of similarity measures using phantom images

This chapter compares the accuracy and robustness of the similarity measures described in chapter 5 when registering a CT volume to a fluoroscopy image of a spine phantom. The chapter begins by describing the spine phantom and then details how a "gold-standard" 2D-3D registration was calculated using fiducial markers. The chapter then describes a number of experiments which assess the accuracy and robustness of the registration algorithm when using different similarity measures. The effect of a number of factors on the accuracy and robustness have been investigated. Experiments have been carried out to determine the effect of overlaying clinical features, such as soft tissue structures and interventional instruments on the phantom fluoroscopy image. Also, the effects of altering CT slice thickness, of altering the out-of-plane translation direction and the effect of including non-physical rays in the DRR have been investigated. In each instance the accuracy of the registrations are compared to the "gold-standard" registration.

6.1 Phantom images with a "gold-standard" registration

6.1.1 The spine phantom

Phantom images were used as it is an extremely difficult task to obtain an accurate "gold-standard" registration using clinical images. The spine phantom is shown in figure 6.1. It consisted of five lumbar vertebrae and a pelvis encased in acrylic which is approximately tissue equivalent at diag-



(a) (b)

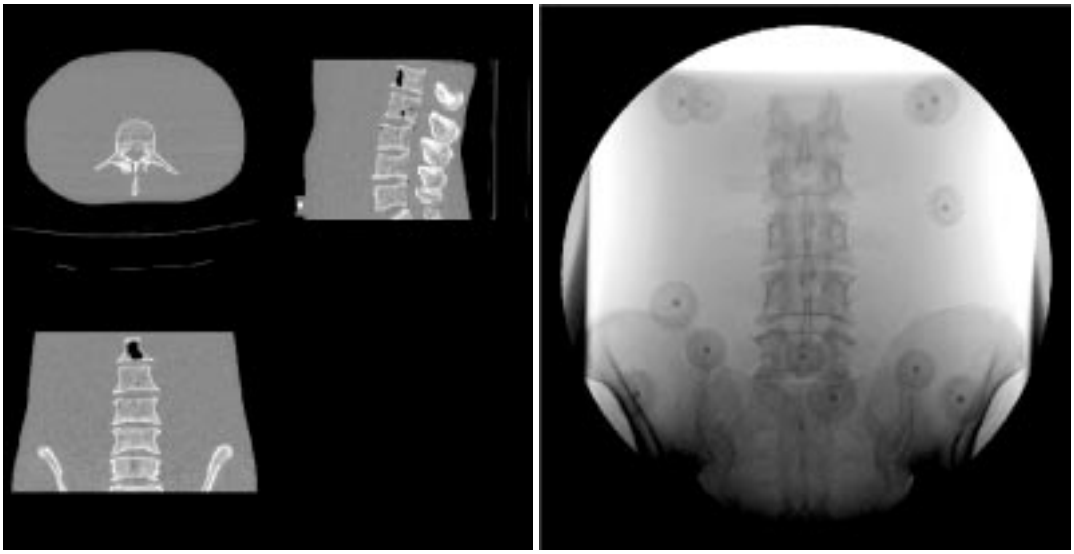
Figure 6.1: Anterior (a) and posterior (b) photographs of spine phantom.

nostic x-ray energies. The acrylic provides the resultant images with some of the characteristics of soft tissue, such as additional attenuation and scatter. However, the acrylic is homogeneous and so does not model the variations in soft tissue which are visible in clinical images.

A CT scan was acquired of the phantom using a Philips TOMOSCAN SR 7000, which had voxel sizes $1.094 \times 1.094 \times 1.5$mm and image dimensions $320 \times 320 \times 119$ voxels, see figure 6.2(a). The same imaging protocol was also used to scan a cubic phantom of known dimensions. The CT volume of the cubic phantom was used to calculate and correct for any geometric scaling errors in the CT volume of the spine phantom using a method described in Hill *et al.* [59]. A fluoroscopy image of the phantom was taken on a Philips multi DIAGNOST 3, which had image dimensions of 1024×1024 pixels and a field of view of 38cm, see figure 6.2(b). A distortion correction phantom was also imaged with the fluoroscopy gantry in the same position. This image was used to distortion correct the fluoroscopy image of the phantom using software supplied by Philips Medical Systems EasyVision Advanced Development which is based on a paper by Haaker *et al.* [53]: see section 4.4.2 for more details.

6.1.2 Fiducial marker design

A "gold-standard" registration between the fluoroscopy and CT images of the spine phantom was calculated using fiducial markers. The fiducial markers were designed so that they could be accurately localised in both the fluoroscopy and CT images. The first requirement was that they should be visible in both modalities and induce no artifacts. Aluminium was chosen as it has a similar linear attenuation coefficient to compact bone at typical fluoroscopy and CT x-ray energies [62]. The next requirement was that the markers must have a particular 3D point which can be firstly,



(a) (b)

Figure 6.2: Reformatted axial (top left), sagittal (top right) and coronal (bottom left) slices from a CT image of a lumbar spine embedded in an acrylic matrix (a). Note one of the fiducial markers is visible in the sagittal slice. Fluoroscopy image of the spine phantom (b). Note the attenuation from the tissue equivalent acrylic in which the spine is embedded and the spherical fiducial markers.

accurately located in the CT volume and secondly, it must also be possible to accurately localise the projection of this point in the 2D fluoroscopy image. To accurately locate the projection of the point it is important that the shape of the perspective projection of the fiducial marker does not depend on the view angle. A sphere is the only shape which fulfils this requirement, therefore, aluminium ball bearings were used. An intensity weighted method was used to locate the centre of the markers in the CT volume. The intensity weighted method requires the markers to be bigger than the largest voxel dimension. The markers should also be surrounded by a second homogeneous medium, to a thickness larger than the voxel dimensions, in order to simplify the partial volume effect. Ball bearings of 5mm diameter were used, which were cast in an acrylic resin. A photograph of one of these markers is shown in figure 6.3.

At least six markers are required to calculate a “gold-standard” 2D-3D registration matrix [48] though more commonly, to average out fiducial localisation error, more markers are used. Initial experiments were carried using 8 markers, however, the resultant “gold-standard” matrix was found to have fairly large errors using the method described in section 6.1.4. Because of this, 12 fiducial markers were used to calculate the “gold-standard” matrix. These markers were rigidly glued to the spine phantom prior to imaging, they can be seen in the fluoroscopy image, figure 6.2(b), and a single marker is also visible in the sagittal slice of the CT scan, figure 6.2(a). The fiducials were placed to achieve a fairly even spread around the centre of the phantom and were also placed so that they did not obscure the vertebrae in a posterior-anterior (PA) or lateral

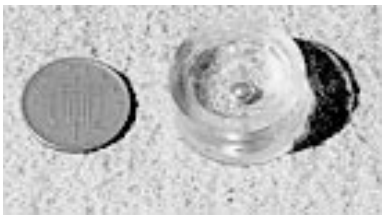


Figure 6.3: Photograph of one of the fiducial markers used to calculated the "gold-standard" registration.

fluoroscopy image. However, one marker did underly the L5 vertebra in the PA fluoroscopy image but this was not found to cause any problems in the subsequent experiments.

6.1.3 Fiducial marker localisation

An intensity weighted method was used to find the centre of the fiducial markers in the CT scan. An initial estimate of the centre of the marker $\mathbf{C}'$, was made by manually picking a point within the marker. This estimate was then refined using equation (6.1) to obtain the centre of the fiducial accurately, $\mathbf{C} = (C_x, C_y, C_z)$

$$\mathbf{C} = \frac{\sum_{i=\mathbf{C}'-d}^{\mathbf{C}'+d} (I_i - I_0)\mathbf{V}_i}{\sum_{i=\mathbf{C}'-d}^{\mathbf{C}'+d} (I_i - I_0)}, \tag{6.1}$$

where d is the diameter of the ball bearing, I_i represents the intensity of the voxel whose centre is at $\mathbf{V}_i = (V_x, V_y, V_z)_i$ and I_0 is the intensity of the background (the acrylic resin). This method is capable of finding the centroid of a marker to sub-voxel accuracy [15, 22]. Wang *et al.* [130] have shown that the above intensity weighted method contains an approximation because it assumes that the contribution from each voxel comes from the centre of the voxel. However, if a fiducial only partially occupies a voxel, then the centroid of that portion of the marker which lies within the voxel is unlikely to coincide with the centroid of the voxel. I believe that the intensity weighted method is suitable for my purposes because, for a sphere of size 5mm, the maximum possible localisation error caused by this approximation is very small, about 0.025mm [130].

Large errors were observed when the intensity weighted method was applied to locate the centre of the fiducials in the fluoroscopy image. This was because of underlying and overlying intensity fluctuations. Lemieux and Jagoe [72] calculated the centre of radio-opaque discs in radiographs using the following method. Starting from a position approximately at the centre of the fiducial, 30 equiangular radial spokes were drawn out. Bilinear interpolation was used to

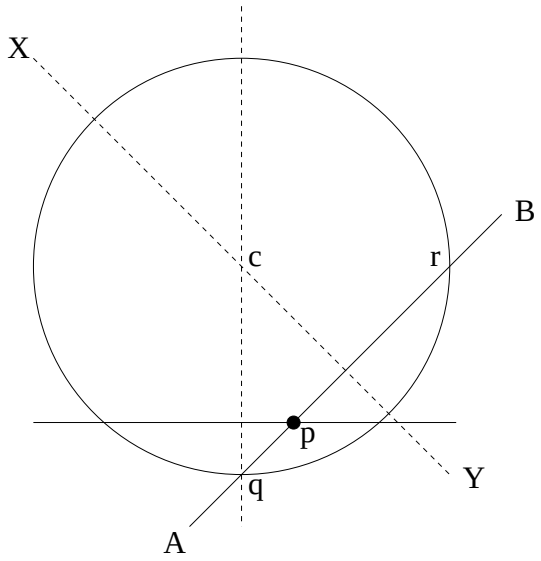
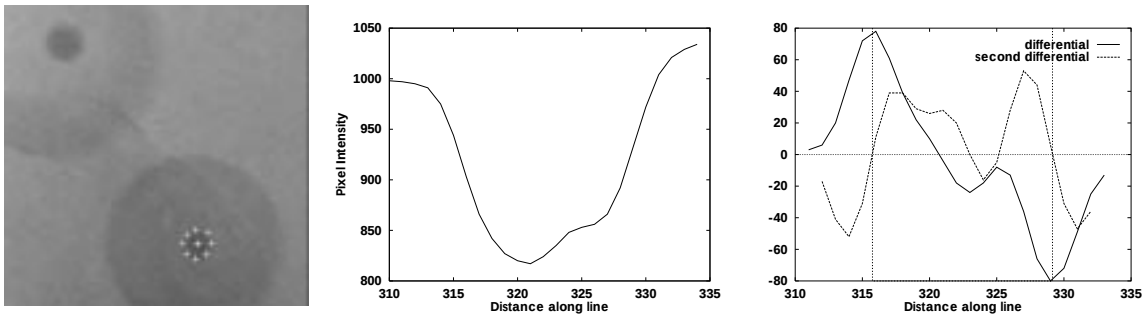


Figure 6.4: How the centre of the fiducial markers were found in the fluoroscopy image. Point p is manually picked and then lines are drawn out at 45^{o} intervals (e.g. line AB). The intercepts between line AB and the edge of marker are found, points q and r. The normal to line AB is drawn equidistant to points q and r to give line XY. Line XY should go through centre c of the fiducial. Triangulation was used to obtain an average value for the position of c.

estimate the pixel intensity along these spokes. The pixel intensities were used to calculate a mean intensity inside and outside the fiducial. These mean values were used to calculate the intercept between each spoke and the edge of the fiducial to sub-pixel accuracy. The centre of the marker was then estimated from these boundary locations.

Lemieux and Jagoe's [72] method was not suitable for my fiducials because they are spherical rather than disc shaped. Instead the following method was used. A pixel within the marker was manually picked (e.g. point p in fig. 6.4) and lines drawn out from this pixel at 45° intervals (e.g. line AB). Each of these lines were taken in turn and a graph plotted of intensity as a function of distance, figure 6.5(b). This plot was differentiated to calculate the positions of steepest upslope and steepest downslope. These positions should correspond to points q and r which are near to the edge of the marker. To obtain sub-pixel accuracy a second order derivative was also calculated to find the positions of the points of inflection. Plots of the first and second order derivatives are shown in figure 6.5(c), where the vertical lines mark the position of the points of inflection, the zero crossings of the second derivative. A normal to line AB (line XY) was then drawn equidistant from points q and r. This normal should go through c, the centre of the marker. Normals are found for each of the four lines at 45° to each other and their intersection points found. The centre of the marker is taken to be the average position of these intersection points. Figure 6.5(a) shows an enlarged section of the fluoroscopy image which contains two fiducial markers. A point has

been picked in the lower fiducial. The white crosses show the positions found around the edge of the fiducial marker and the position of the centre of the fiducial marker calculated using the edge positions.



(a) (b) (c)

Figure 6.5: How the fiducial markers are localised in the fluoroscopy image. Image (a) shows an enlarged region from a fluoroscopy image which contains two fiducial markers. A point has been picked in the lower marker and the white crosses show the eight positions found around the edge of the marker and the centre of the marker calculated from these edge positions. The pixel intensities along a line through the marker (e.g. line AB in figure 6.4) are shown in (b). The first and second differential of the graph shown in (b) are shown in (c). The vertical lines mark the zero crossing of the second differential which is used to find the max and min positions of the first differential to sub-voxel accuracy.

6.1.4 Calculation of "gold-standard" accuracy

Once the position of the fiducial markers had been found, a "gold-standard" registration matrix was calculated using the SVD method described in section 4.2.1. This matrix was then decomposed into the ten perspective projection parameters, $\mathbf{W}$, using the method described in section 4.2.2.

I have used the same notation as in Maurer *et al.*[87] to describe the errors in the "gold-standard" registration matrix. In Maurer *et al.*[87] a number of different error measures are defined. Target registration error (TRE) is defined as the error between two "gold-standard" target points which were not used to calculate the registration matrix, see figure 6.6. The calculation of TRE involves the localisation of the target points which will have an associated error involved, target localisation error (TLE). Observed target registration error (TRE_o) is the measured value which is combination of the true target localisation error (TRE_t) and the TLE.

Observed target registration error (TRE_o) of the "gold-standard" 2D-3D registration was calculated in the following way. Two fiducial markers were removed from the set of twelve so they could be used as target points. A matrix was calculated using the remaining markers and used to project the 3D positions of the two target points into 2D. The distance between the projected 2D

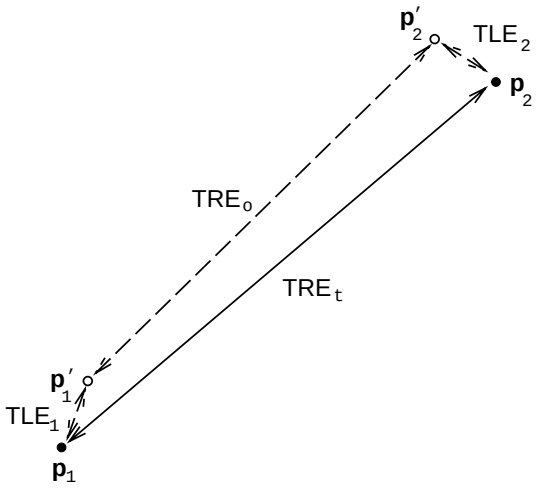


Figure 6.6: The relationship between true target registration error (TRE_t), observed target registration error (TRE_o) and target localisation error (TLE). The points $\mathbf{p}_1$ and $\mathbf{p}_2$ are the "true" position of two fiducial markers which were not used to calculate the "gold-standard" registration matrix. $\text{TRE}_t = \|\mathbf{p}_1 - \mathbf{p}_2\|$. Due to errors in calculating the position of $\mathbf{p}_1$ and $\mathbf{p}_2$ (TLE_1 and TLE_2) their measured position is actually $\mathbf{p}'_1$ and $\mathbf{p}'_2$. $\text{TRE}_t = \|\mathbf{p}'_1 - \mathbf{p}'_2\|$. If TLE_1 and TLE_2 are uncorrelated then $(\text{TRE}_o)^2 \approx (\text{TRE}_t)^2 + 2(\text{TLE})^2$ [87].

position and the "gold-standard" 2D point position was calculated. Distances were calculated for each possible set of ten points from twelve, resulting in 132 distances. The RMS value of these distances was 1.2 pixels.

In my registration evaluation, errors are calculated for each of the extrinsic parameters. It was, therefore, important to be able to relate the 2D error, TRE_o, to errors in each of the extrinsic parameters. This was achieved by carrying out the following numerical simulation. The "gold-standard" 3D points were projected into 2D to form a "perfect" set of 3D and 2D points. Two target points were removed from the set of twelve and a random Gaussian distributed error (standard deviation ε pixels) was added to the 2D positions of the remaining 10 points to simulate true target registration error (TRE_t) [87]. A new perspective projection matrix was calculated using an iterative scheme which altered the extrinsic parameters to minimise the RMS distance between the 3D points projected into 2D, and the 2D points with random error added. The intrinsic parameters, $\mathbf{I} = (k_1, k_2, c_s, l_s)$, were kept fixed at their "gold-standard" values. The RMS error between the target points projected into 2D using the new projection matrix and their "perfect" 2D positions was calculated. This RMS error represents the TRE_t of the new matrix and hence also the TRE_t of the new set of extrinsic parameters. This process was repeated 15 times for each possible set of ten points from twelve resulting in 990 values of TRE_t, from which an average was calculated.

	rigid body parameters					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	δx	δy	δz
SD	0.09	0.09	0.09	1.34	0.15	0.15

Table 6.1: Expected standard deviation of "gold-standard" parameters calculated using 990 sets of parameters which, when used to project 3D points to 2D, give an average RMS error of 1.2 pixels in the position of the 2D points.

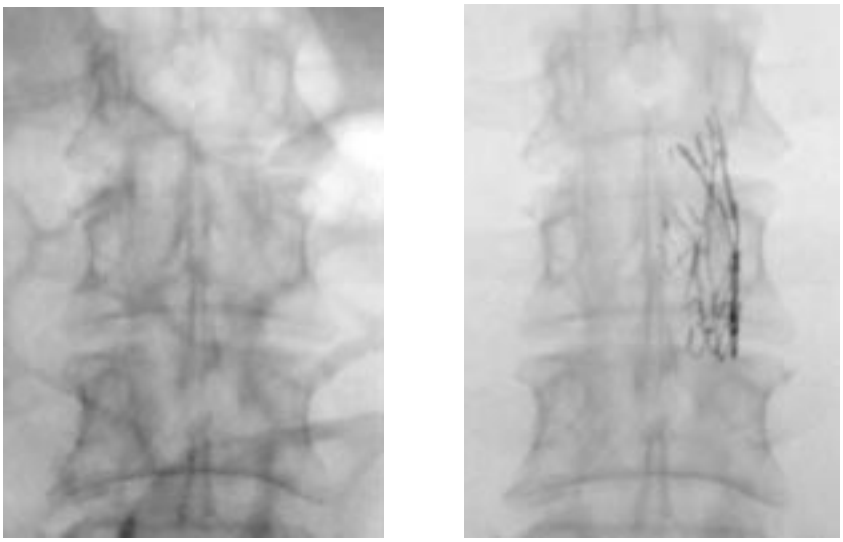
By experiment it was found that $\varepsilon = 1.35$ pixels produced an average TRE_t of 1.2 pixels, which is equal to the TRE_o of the "gold-standard" matrix. Therefore, this experiment created 990 sets of extrinsic parameters which, when used to project points from 3D to 2D, have an average TRE_t of 1.2 pixels. The standard deviations of these 990 sets of parameters are shown in table 6.1. They give a good indication of the expected error in each of the "gold-standard" extrinsic parameters.

The TRE_o is made up from TRE_t and a target localisation error, TLE. If TLE is the same in each image space and uncorrelated then $(\text{TRE}_o)^2 \approx (\text{TRE}_t)^2 + 2(\text{TLE})^2$ [87]. In the above TLE has been assumed to equal zero and so an upper bound for both TRE_t and for the expected errors in each of the extrinsic parameters has been calculated.

6.1.5 Simulating clinical images

The fluoroscopy image of the phantom was enhanced by overlaying soft tissue structures and an interventional stent, so that the effect of these clinical features on the performance of the similarity measures could be investigated. Three clinical images were simulated by adding soft tissue, an interventional stent, and both soft tissue and a stent. No soft tissue structures were introduced to the CT volume as these would be removed during the thresholding stage, see section 4.3.3.

Overlaying soft tissue – A region that contained a variety of soft tissue structures was manually segmented from a clinical fluoroscopy image acquired using a Philips MD3 fluoroscopy set at 38cm magnification setting. The major soft tissue structure visible was the small bowel, which produced some large variations in pixel intensity due to the presence of gas. The segmented soft tissue was overlaid on top of the original spine phantom fluoroscopy image by adding the pixel intensities of both images together, using a weighting factor ($w = 0.7$), $I_{combined} = I_{phantom} + wI_{softtissue}$, the resultant image is shown in figure 6.7(a). The weighting factor was necessary to achieve a visually realistic ratio between the contributions from bone and soft tissue. It was chosen by visual inspection: comparing the simulated image with clinical images.



(a) (b)

Figure 6.7: Enlarged portion of the fluoroscopy image, (a) with soft tissue structures overlaid and (b) with interventional stent overlaid over the L3 vertebra.

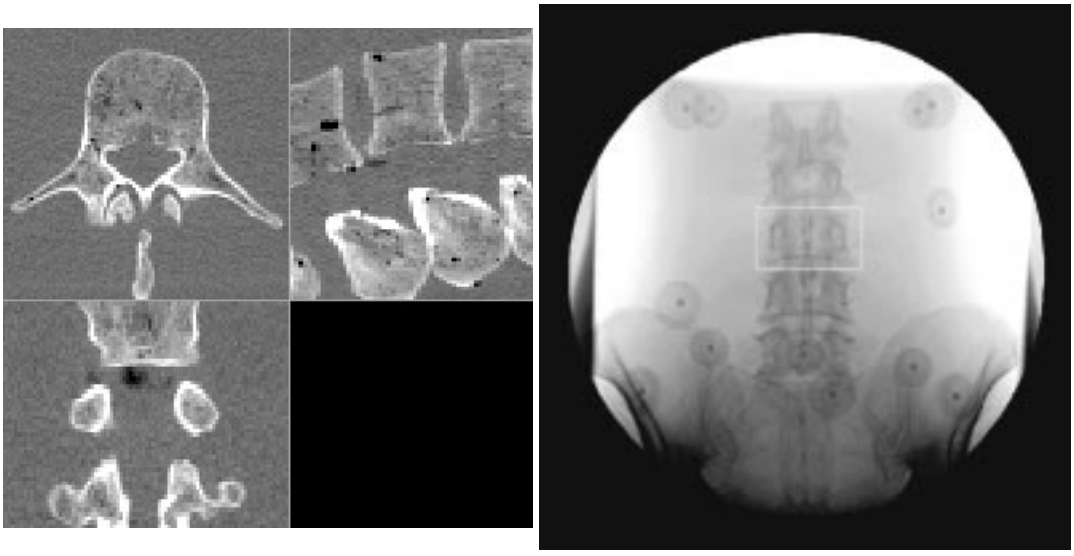
An interventional stent – An image of part of an interventional stent was segmented from a clinical fluoroscopy image. The image was again acquired using a Philips MD3 fluoroscopy set at 38cm magnification setting. In the chosen image a large section of the stent did not overlie any bony structures. This section of the stent was manually segmented. The stent was then separated from the overlying and underlying soft tissue structures by thresholding. The thresholded image was overlaid on top of the spine phantom fluoroscopy image using the weighting factor, $w = 1.0$. The resultant image is shown in figure 6.7(b). The objective was to create an overlap with a large number of sharp edges which is characteristic of a number of interventional instruments e.g. stents, guidewires or needles.

The third image was a combination of the above two images with both soft tissue and a stent overlaid.

Figure 6.7 shows the interventional stent and soft tissue structures overlying the L3 vertebra. Similar images to those shown in figure 6.7 were also produced with the soft tissue and stent overlying the L1, L2, L4 and L5 vertebrae.

6.2 Accuracy and robustness of similarity measures

An experiment was carried out to investigate how accurately and robustly seven similarity measures registered a CT scan to a fluoroscopy image of the spine phantom. Registrations were carried



(a) (b)

Figure 6.8: Reformatted axial, sagittal and coronal slices through the reduced CT volume surrounding the L3 vertebra (a). Fluoroscopy image showing region of interest (white rectangle) used for registrations to the L3 vertebra (b).

out to each of the lumbar vertebrae. The fluoroscopy image was blurred using an ellipse of minor and major axes 3x5 pixels, as explained in section 4.4.1. For each vertebra a reduced CT volume (with the vertebra of interest approximately in the centre) was segmented from the CT image and a ROI was defined in the fluoroscopy image. Figure 6.8 shows the reduced CT volume and ROI used for registrations based on the L3 vertebra. The transverse processes were intentionally excluded from the fluoroscopy ROI as they are not usually visible in clinical fluoroscopy images. Sixty four registrations were carried out to each vertebra, resulting in 320 registrations for each similarity measure.

6.2.1 Obtaining realistic starting positions for the algorithm.

The starting estimates for registration were the "gold-standard" value plus or minus $\Delta\mathbf{E}$, see table 6.2. There are sixty four possible combinations of "gold-standard" plus or minus $\Delta\mathbf{E}$ which are at the corners of a six dimensional hyper-cuboid. These values were chosen to simulate how the algorithm could operate when registering clinical images which have no "gold-standard" registration. In clinical practice a starting estimate could be found using knowledge of patient position, imaging geometry and by picking corresponding anatomical points in the images. It is much easier to obtain an accurate estimate of some parameters, such as the in-plane translations (Y and Z) compared to other parameters, especially the out-of-plane translation (X).

The values of $\Delta\mathbf{E}$ are two times the estimated accuracy of a 2D-3D registration carried

	Displacement of extrinsic parameters					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
$\Delta\mathbf{E}$	3.4	7.6	7.8	50.8	3.6	2.4

Table 6.2: The displacement ($\Delta\mathbf{E}$) of the starting positions from "gold-standard" for each of the six rigid body parameters.

out by manually picking corresponding anatomical points in the two images. These values were calculated using the following experiment. Six anatomical points were picked on the spine and projected into 2D using the "gold-standard" matrix. A random Gaussian distributed error was added to the position of these points to simulate observer error. The size of the 3D error (3mm standard deviation) was based on the largest RMS error found in a reproducibility experiment to find landmarks in a CT volume described in Hill *et al.* [57]. A similar experiment for 2D errors was not found in the literature and so the 2D error was roughly estimated to have a standard deviation of 6 pixels. An iterative scheme was used to alter the extrinsic parameters to minimise the RMS distance between the new 3D points projected into 2D and the new 2D points. The intrinsic parameters were kept fixed at their "gold-standard" values. This was carried out 150 times and a mean and standard deviation of the final extrinsic parameters calculated. The values of $\Delta\mathbf{E}$ were set to two times this standard deviation. It is very unlikely for all the starting positions to be out by two times the standard deviation (0.000002% chance) and so this represents a rigorous test for the algorithm.

It is important for the registration algorithm that, at the starting position some corresponding features are at least partially overlapping. One method of achieving this is to pick a single corresponding point in the DRR and in the fluoroscopy image and overlay these points. The starting positions for registration were modified by simulating this single point picking exercise. The "gold-standard" matrix was used to calculate a corresponding 2D and 3D point at the top of the spinous process. Before the registration began, the 3D point was cast into 2D using a matrix calculated from the initial parameters. The position of this 2D point and the "gold-standard" 2D point were both altered by adding a random Gaussian distributed error (standard deviation 6 pixels) to simulate observer error. These two points were then overlaid by altering the in-plane translational parameters (Y and Z). Figure 6.9 shows three DRRs of the L3 vertebra produced at typical starting positions and the L3 vertebra from the fluoroscopy image for reference.

To summarise, the starting position for my algorithm was the "gold-standard" position plus or minus twice the expected accuracy of a 2D-3D registration algorithm based on manually pick-

ing anatomical points. However, the in-plane parameters were also altered by overlaying a single 2D point from the DRR and the fluoroscopy image. This latter stage was necessary so that corresponding features were at least partially overlapping at the registration starting position.

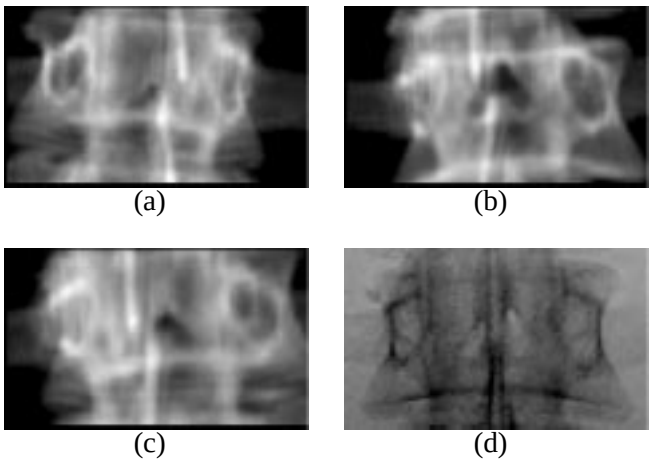


Figure 6.9: Three DRRs (a),(b) and (c) produced at typical starting positions i.e. the "gold-standard" value plus or minus $\Delta\mathbf{E}$ (table 6.2). The corresponding region of the fluoroscopy image is shown for comparison in (d).

6.2.2 Registration evaluation

For each registration the algorithm provides a final set of extrinsic parameters, $\mathbf{E}$. From a given experiment the algorithm provides a number of final registration positions $\mathbf{E}_i, i = 1 \rightarrow N$, where N is the total number of registrations attempted (64 per vertebra in these experiments). These parameters have been compared with the "gold-standard" extrinsic parameters $\mathbf{E}_{GS}$, in the following ways. Firstly, failed registrations were removed from the set of final registration positions. A final registration position was defined as a failure if any of the extrinsic parameters were further from "gold-standard" than $\Delta\mathbf{E}$ as shown in table 6.2, i.e. if

$$\|E_i^k - E_{GS}^k\| > \Delta E^k, \qquad k = 1 \rightarrow 6, \tag{6.2}$$

where the k denotes the different elements which make up the set of extrinsic parameters $\mathbf{E} = (\theta_x, \theta_y, \theta_z, X, Y, Z)$, i.e. $E^1 = \theta_x, E^2 = \theta_y, \ldots, E^6 = Z$. Similarly E_{GS}^k and ΔE^k represent the different elements in $\mathbf{E}_{GS}$ and $\Delta\mathbf{E}$. This leaves a number, N_{good}, of successful registrations.

A RMS error, $\mathbf{RMSE} = (\theta_{x_{RMSE}}, \theta_{y_{RMSE}}, \theta_{z_{RMSE}}, X_{RMSE}, Y_{RMSE}, Z_{RMSE})$, was calculated from the remaining final registration positions ($i = 1 \rightarrow N_{good}$).

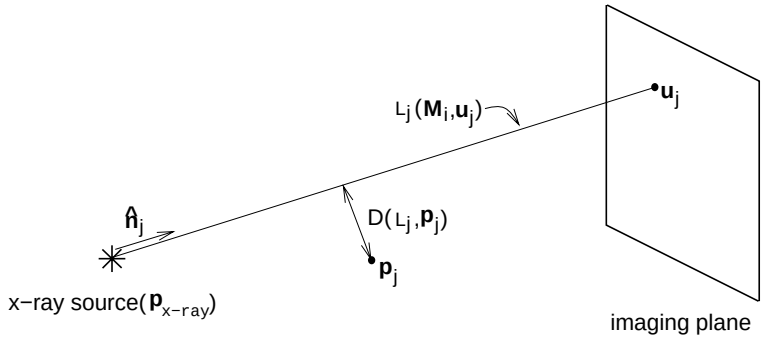


Figure 6.10: Reprojection distance, as defined by Masutani *et al.* [85] is the average value of $D(L_j, \mathbf{p}_j)$ calculated for a number of fiducial markers, $j = 1, \ldots, N_{fiducials}$. Where $L_j(\mathbf{M}_i, \mathbf{u}_j)$ represents a 3D line from the 2D position of the j'th fiducial marker ($\mathbf{u}_j$) back to the x-ray source, calculated using the final perspective projection matrix for the i'th registration, $\mathbf{M}_i$. $D(L_j, \mathbf{p}_j)$ is the minimum distance between line L_j and the 3D position of the j'th fiducial marker, $\mathbf{p}_j$.

$$RMSE^k = \sqrt{\frac{1}{N_{good}} \sum_{i=1}^{N_{good}} (E_i^k - E_{GS}^k)^2}, \qquad k = 1 \rightarrow 6 \tag{6.3}$$

By calculating $\mathbf{RMSE}$ it is possible to investigate which of the extrinsic parameters are the most and least accurate. However, because six error values are obtained, rather than a single measure of accuracy, it can be difficult to relate the values in $\mathbf{RMSE}$ to the final registration accuracy when transferring positions from 3D to 2D or vice-versa. Most of the registration techniques reviewed in chapter 3 use two 2D images and calculate a 3D error in mm [51, 70, 73, 77]. Brown *et al.* [49] also uses a 3D error in mm even though their algorithm registers to a single 2D image. When registering to a single 2D image the largest errors usually occur along the source-to-film direction. However, if information is to be projected from 3D to 2D, then fairly large errors can be tolerated along the source-to-film axis without causing large errors in the position of projected information. Because of this, reprojection distance has been chosen to give a single value for the registration errors in the experiments described in this chapter.

$$d_i' = \frac{1}{N_{fiducials}} \sum_{j=1}^{N_{fiducials}} D(L_j(\mathbf{M}_i, \mathbf{u}_j), \mathbf{p}_j) \tag{6.4}$$

Reprojection distance, d', as defined by Masutani *et al.* [85], is given by equation (6.4) and is explained in figure 6.10. The reprojection distance is the average value of $D(L_j, \mathbf{p}_j)$ calculated over a number of fiducial markers. Where $D(L_j, \mathbf{p}_j)$ represents the minimum distance between a line L_j and the 3D position of the j'th fiducial marker $\mathbf{p}_j$. The line L_j is drawn from the 2D

position of the j'th fiducial marker $\mathbf{u}_j$ to the x-ray source p_{x-ray}. This line is calculated using the final perspective projection matrix for the i'th registration, $\mathbf{M}_i$. The minimum distance D between a 3D point $\mathbf{p}_j$ and a line $L_j = p_{x-ray} + \lambda\hat{\mathbf{n}}_j$ is given by equation 6.5. Where the line L_j goes through point p_{x-ray}, is parallel to the unit vector $\hat{\mathbf{n}}_j$ and λ is an arbitrary scalar (see figure 6.10).

$$D(L_j, \mathbf{p}_j) = \|(\mathbf{p}_j - \mathbf{p}_{x-ray})\times\hat{\mathbf{n}}_j\| \tag{6.5}$$

The value of d' from a given set of fiducials depends not only on the final registration error, but also on the configuration of the fiducials. If the fiducials are a long distance from the axes of rotation then the value of d' will be larger than if the fiducials were close to the axes of rotation (unless the error is purely translational).

In this thesis reprojection distance, d, is calculated over all the voxels in the CT volume which project onto the fluoroscopy screen, rather than over a number of fiducial markers, see equation (6.6). The "gold-standard" 3D positions, $\mathbf{p}_j$, are the centre of the voxels. The "gold-standard" 2D positions, $\mathbf{u}_j$ are the centre of the voxels projected into 2D using the "gold-standard" registration matrix, i.e. $\mathbf{u}_j = \mathbf{M}_{GS}\mathbf{p}_j$. Unlike Masutani *et al.* [85], this measure will not be affected by the configuration of fiducial markers. However, it will be affected by the dimensions of the CT volume. As the size of the CT volume increases then, on average, so will the size of d. However, as we are only interested in projecting information from within the CT volume, then d can be said to represent the average error of a point projected from within the CT volume.

$$d_i = \frac{1}{N_{voxels}} \sum_{j=1}^{N_{voxels}} D(L_j(\mathbf{M}_i, \mathbf{M}_{GS}\mathbf{p}_j), \mathbf{p}_j) \tag{6.6}$$

Knowledge of the average reprojection distance per voxel might also be useful in a clinical situation. For example, if an interventionist wanted to avoid a critical structure, such as a nerve or blood vessel, it could be segmented from the CT scan and dilated by d before projecting. This would result in the structure plus an average error boundary being projected into the 2D image. Similarly, if there is a structure which the interventionist is aiming for, such as a metastasis, then by eroding the structure by d the resultant projection will be reduced by the average error. If the structures are also projected without dilation or erosion then this should give the interventionist a good indication of the expected error on the fluoroscopy screen, see figure 6.11.

In the following and subsequent results sections, reprojection distance is calculated and averaged over a number of registrations to give $\overline{d}$.

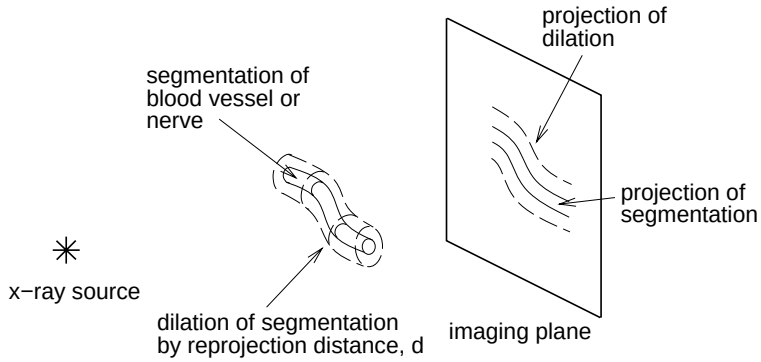


Figure 6.11: How errors might be shown in a clinical situation. If a structure which an interventionist wants to avoid, such as a nerve or blood vessel, is dilated by the reprojection distance, d, it is possible to project the structure and an average error boundary.

$$\overline{d} = \frac{1}{N_{good}} \sum_{i=0}^{N_{good}} d_i \tag{6.7}$$

To show the accuracy of the registration within the region of interest $\overline{d(ROI)}$ is also calculated. This represents the reprojection distance calculated using all the voxels within the reduced CT volume which project to the fluoroscopy ROI.

By using the 990 sets of $\mathbf{E}$ which create an average error of 1.2 pixels on the fluoroscopy screen (see section 6.1.4) it was possible to estimate the values of $\overline{d}$ and $\overline{d(ROI)}$ for the "gold-standard" registration. The values of $\overline{d}$ and $\overline{d(ROI)}$ were also estimated for a set of 64 typical starting positions (see section 6.2.1). These values are shown in table 6.3.

	$\overline{d}$ (mm)	$\overline{d(ROI)}$ (mm)
Errors in "gold-standard" registration	0.36	0.22
Errors at typical starting positions	20.2	5.3

Table 6.3: Estimates of average reprojection error for the "gold-standard" registration and for a set of 64 typical starting positions.

6.2.3 Results

Results are presented in the following tables. They show $\mathbf{RMSE}$, $\overline{d}$, $\overline{d(ROI)}$ and the failure rate. The average registration took 74 seconds on a Sun Ultra 30 (300 MHz).

Table 6.4 shows the results of registering to each of the lumbar vertebra in the fluoroscopy image with no additional structures overlaid. It was noted while carrying out these experiments that a large number of failed registrations and large RMS errors occurred when some of the simi-

Similarity measure	Type	RMSE rotation (degrees) θ_x	θ_y	θ_z	RMSE translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d_{(ROI)}}$ (mm)	No. fail, (%)
Cross correlation	2 B	0.57	1.61	0.43	9.6	0.72	1.10	2.46	0.92	19.1
Entropy	2 A	0.42	0.92	0.43	10.0	0.70	1.10	2.03	0.80	14.4
Mutual Information	2 A	0.67	1.33	0.88	19.4	0.97	1.13	3.02	1.03	36.6
Gradient correlation	3 B	0.25	0.33	0.27	5.4	0.44	0.58	1.11	0.51	0.0
Pattern Intensity	1 A	0.30	0.35	0.31	5.8	0.51	0.49	1.18	0.54	0.0
Gradient difference	3 A	0.28	0.27	0.32	6.3	0.46	0.53	1.14	0.52	0.9
Local correlation	1 A	0.25	0.41	0.38	4.9	0.58	0.51	1.23	0.58	1.3

Table 6.4: The performance of the similarity measures based on 320 registrations (64 to each lumbar vertebra) to the fluoroscopy image of spine phantom.

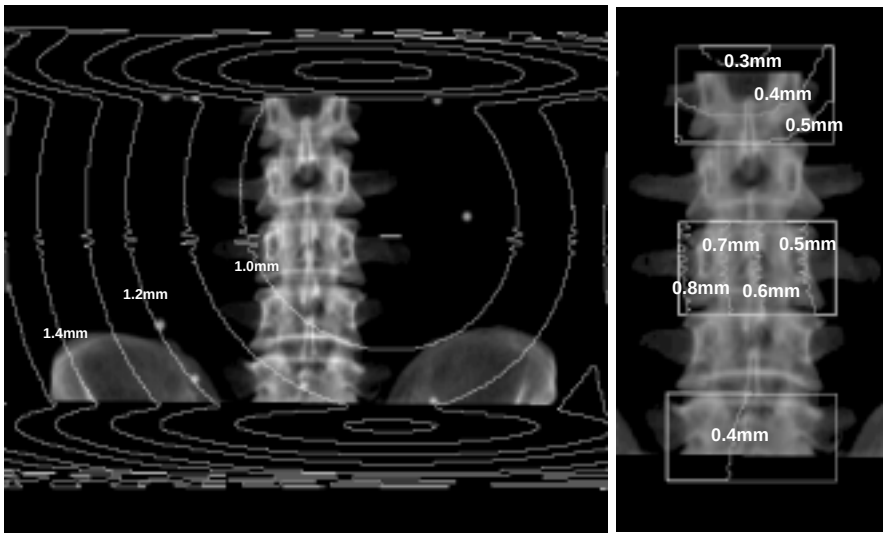
larity measures registered to the L1 vertebra. This could be due to a number of reasons. Firstly, the vertebra is near the top of the fluoroscopy image where intensity fluctuations due to the heel effect are strongest. Secondly, the phantom does not have any thoracic vertebrae and so the L1 vertebra does not have a vertebra directly above it. Due to the interleaved nature of vertebrae, all the regions of interest apart from the one around the L1 vertebra will possess some bony structures from the vertebra above. Therefore, there is less bony information in the L1 ROI on which to register. Thirdly, the phantom only has a few millimetres of acrylic above the L1 vertebra. Because of this the fluoroscopy image near the top of the phantom may be unrealistic, as rays passing through the L1 vertebra may also exit out from the top of the phantom. This physical truncation effect could cause varying contributions from the acrylic which would not be modelled in the DRR.

Due to the problems with the L1 vertebra, some of which may be caused by the physical structure of the phantom rather than the performance of the similarity measures, it was decided to remove this vertebra from the experiments and only register on vertebrae L2 — L5. Table 6.5 shows the results of registering on these four vertebrae.

The results from table 6.5 show that mutual information was the least accurate similarity measure, with $\overline{d}$ = 2.82mm and a 25% failure rate. The measures with the next largest errors were cross correlation and entropy, with values of $\overline{d}$ approximately equal to 2mm and a failure rate around 4%. The most accurate measures were local correlation, pattern intensity, gradient difference and gradient correlation which had errors, as measured by $\overline{d}$, of just over 1mm and hardly any failures, only 2 from 1,000 registrations.

Similarity measure	Type	RMSE rotation (degrees) θ_x	θ_y	θ_z	RMSE translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d(ROI)}$ (mm)	No. fail, (%)
Cross correlation	2 B	0.42	0.87	0.40	9.7	0.65	0.90	2.06	0.83	3.5
Entropy	2 A	0.35	0.65	0.41	8.8	0.68	0.82	1.81	0.73	5.1
Mutual Information	2 A	0.52	1.30	0.79	18.4	0.66	0.77	2.82	0.96	25.0
Gradient correlation	3 B	0.22	0.34	0.28	4.6	0.44	0.41	1.02	0.53	0.0
Pattern Intensity	1 A	0.26	0.37	0.32	4.8	0.51	0.39	1.10	0.57	0.0
Gradient difference	3 A	0.24	0.29	0.32	5.4	0.46	0.43	1.06	0.56	0.0
Local correlation	1 A	0.25	0.32	0.37	4.0	0.56	0.50	1.14	0.58	0.8

Table 6.5: The performance of the similarity measures based on 256 registrations (64 to each vertebra; L2 — L5) to the fluoroscopy image of spine phantom.



(a) (b)

Figure 6.12: Contoured plots of $\overline{d(u,v)}$ over a DRR produced using the "gold-standard" parameters; (a) shows $\overline{d(u,v)}$ over the entire CT volume and DRR; (b) shows $\overline{d(u,v)}$ over the reduced CT volumes and ROI for the L1, L3 and L5 vertebrae. The outer contour in (b) is the limit of the fields of view drawn around each vertebra.

Figure 6.12 shows how the contribution to the overall error measures, $\overline{d}$ and $\overline{d(ROI)}$, varies over the 2D field of view. The contours show the value of $\overline{d(u,v)}$, see equation (6.8), which represents the average projection error along a ray path from the source to pixel (u,v) where $N_{voxels}(u,v)$ represents the number of voxels along the ray path: see equations (6.5) and (6.6) and figure 6.10 for more information on the arguments of parameter D. The registration results using the gradient correlation similarity measure were used to produce figure 6.12. This measure was chosen as it was the most accurate measure according to table 6.4. For figure 6.12(a) the whole CT volume and the results from registrations to all five vertebrae were used. Figure 6.12(b) shows the results using the reduced CT volumes, ROI and registration results for the L1,L3 and L5 vertebrae.

$$\overline{d(u,v)} = \frac{1}{N_{good}} \sum_{i=0}^{N_{good}} \frac{1}{N_{voxels}(u,v)_i} \sum_{j=1}^{N_{voxels}(u,v)_i} D(L_j(\mathbf{M}_i, \mathbf{M}_{GS}\mathbf{p}_j), \mathbf{p}_j) \qquad (6.8)$$

Figure 6.12 (a) shows that the registration is accurate to within 1.2 mm over all of the vertebrae. Truncated rays have been removed from the DRR in figure 6.12 to show how the behaviour of $\overline{d(u,v)}$ changes in the truncated region. This effect is caused because the truncated rays do not traverse the entire length of the CT volume. The average distance of rays to the centre of rotation, and hence also the effect of rotational errors, will therefore alter in a different fashion for truncated rays compared to non-truncated rays.

The values of $\overline{d(u,v)}$ have been calculated for the truncated rays for the following reason. Although truncated rays are not included in the registration process (as they have no physical counterpart in the fluoroscopy image), there is a difference between producing a DRR for registration purposes and projecting information from the CT volume once the two images have been registered. Once registration has occurred whether a ray is truncated or not is no longer of interest. The registration is used to transfer information from one image to another. Therefore, for example, if an important feature such as a nerve was present near to the top of the CT volume, it does not matter whether or not the feature projects into the truncated region or the non-truncated region of the DRR, the location of the projected feature on the fluoroscopy screen will be just as valid and so the associated error value $\overline{d(u,v)}$ is of interest.

Figure 6.12 (b) shows the expected accuracy when projecting information to the ROI around the L1,L3 and L5 vertebrae. The outer contour shows the limits of the ROIs used during registration. The plot shows high accuracy, 0.3 – 0.4 mm, in some areas and the majority of the plot is accurate to 0.6mm.

The least accurate of the seven similarity measures was mutual information. Mutual infor-

mation has shown itself to be very useful in 3D multi-modality registration [81, 114, 115, 124]. However, when registering 3D images, there are usually a large number of voxels to populate the 2D histogram, around 2 million in an image of size 256×256×30 voxels. In the case of 2D-3D registration there are far fewer pixels to populate the 2D histogram. For example there are only 390 pixels within the ROI of the L3 vertebra at the starting resolution of 128×128 pixels. Therefore, mutual information may be failing as the 2D histogram is relatively sparsely populated having an average of only 0.4 pixels per bin, compared to the 1D histogram used by entropy of the difference image which has an average of 6 pixels per bin at the starting resolution. This was investigated by carrying out a further experiment. A larger ROI was drawn in the fluoroscopy image so that the L2, L3 and L4 vertebrae were included, effectively trebling the number of pixels used in the registration. A reduced CT volume was also created which contained the L2, L3 and L4 vertebrae. Sixty four registrations were carried out using mutual information, the larger ROI and the new reduced CT volume. The results of this experiment are shown in table 6.6.

Similarity measure	Type	**RMSE** rotation (degrees)			translation (mm)			$\overline{d}$ (mm)	$\overline{d_{(ROI)}}$ (mm)	No. fail, (%)
		θ_x	θ_y	θ_z	X	Y	Z			
Mutual Information	2 A	0.22	0.71	0.33	4.3	0.41	0.43	1.33	0.68	0.0

Table 6.6: Registration results for the mutual information similarity measure based on 64 registrations to a region of interest in the fluoroscopy image containing the L2, L3 and L4 vertebrae.

6.2.4 Discussion

By comparing the results using all five vertebrae (table 6.4) with the results for registering on vertebrae L2 — L5 (table 6.5) it is clear that cross correlation, entropy and mutual information failed a large number of times and had large RMS errors when registering to the L1 vertebra. All these measures are type '2', see table 5.1.

Type '2' measures register using pixel intensities and their region of comparison is the whole ROI or area of overlap within the ROI. The three reasons why registrations to the L1 vertebra are expected to be more difficult than registrations to the other vertebrae were discussed in the previous section. Two of these reasons, the increased influence of the heel effect and truncation effects caused by the shape of the phantom, would create underlying and overlying intensity fluctuations in the fluoroscopy image. Type '1' measures are expected to be able to filter out these effects if they are on a larger scale than the local region used by these measures (e.g. larger than the radius r). Type '3' measures are also expected to be able to filter out these effects as they register using

gradient images and the above mentioned heel and truncation effects are not expected to introduce sharp edges into the fluoroscopy image. Type '2' measures on the other hand have no special features to allow them to filter out these effects, which may be the reason for their high failure rate and large errors when registering to the L1 vertebra.

The results from registering on vertebrae L2 — L5 (table 6.5) show that all the similarity measures, except for mutual information, performed reasonably well. The largest failure rate for these six measures was 5.1% and all of the RMS errors, apart from in the source to film direction, were smaller than 0.9 mm or degrees. The out-of-plane translational errors for these six measures were all less than 9.7mm and the most accurate measure had an RMS error of just 4mm. Achieving this accuracy in the out-of-plane translation direction is very impressive. The difference between two DRRs, one produced at the "gold-standard" position and one at the "gold-standard" position plus a 4mm translation along the source to film direction would be almost indiscernible to the eye. Analysing the values of $\overline{d}$ and $\overline{d(ROI)}$ from table 6.5 shows a big difference in the registration accuracy between type '2' measures and types '1' or '3'. On average the type '2' measures had values of $\overline{d}$ and $\overline{d(ROI)}$ roughly double those of the types '1' or '3'.

Table 6.6 shows the result of registering with the mutual information similarity measure using a larger ROI which included the L2, L3 and L4 vertebrae. Comparing the results from this experiment with those from the earlier experiments (table 6.5) shows that the mutual information measure is much more accurate and robust if the larger ROI is used. The values for $\overline{d}$ and $\overline{d(ROI)}$ from table 6.6 show that mutual information using the large ROI performed better than the cross correlation and entropy measures using a ROI containing a single vertebra. However, the gradient correlation, pattern intensity, gradient difference and local correlation measures were still more accurate than mutual information. These results appear to back up my earlier hypothesis that mutual information is failing because the 2D histogram is too sparsely populated to provide an accurate and robust statistical measure. However, I have only tested this hypothesis using a single experiment. A full investigation into how mutual information performs when the joint histogram is sparsely populated is considered outside the scope of this thesis.

Another possible reason for the poor performance of mutual information is that it does not assume a linear relationship between the intensities in the two images. The DRR has been produced so that this relationship should be close to linear. Measures that assume a linear relationship make implicit use of this extra information, which may account for their greater accuracy.

The statistical significance between the values of $\overline{d}$ and $\overline{d(ROI)}$ obtained using different similarity measures have not been calculated for the following reason. A t-test was carried out to

detect whether or not the difference between the mean values of $\overline{d(ROI)}$ in table 6.5 were significant. The result was that all of the differences were significant ($p>0.005$). This however is an overinterpretation of the results. A calculation of the statistical significance between the values of $\overline{d}$ and $\overline{d(ROI)}$ should also include the errors in the "gold-standard" registration. The best way of achieving this would be to calculate the "gold-standard" registration a number of times using different fluoroscopy images, either different views of the same phantom or images of a different phantom, and repeat the registration experiments. Repeated calculations of the "gold-standard" registration using the same fluoroscopy image would produce very similar registrations due to the semi-automated methods of fiducial location. This is one limitation of the work described in this thesis, only one phantom fluoroscopy image was used. Therefore, variations caused by using different fluoroscopy images and variations in the calculation of the "gold-standard" registration were not investigated.

6.3 Effect of clinical features on the performance of the similarity measures

To investigate how the accuracy and robustness of the similarity measures are affected by clinical features the experiment described in section 6.2 was repeated, but the fluoroscopy image was replaced by one of the simulated fluoroscopy images, see section 6.1.5. Registrations were carried out on vertebrae L2 — L5.

6.3.1 Results

The following tables show $\mathbf{RMSE}$, $\overline{d}$, $\overline{d(ROI)}$ and the number of failures when registering to the fluoroscopy image with soft tissue overlaid (table 6.7), the fluoroscopy image with an interventional stent overlaid (table 6.8) and the fluoroscopy image with both soft tissue and an interventional stent overlaid (table 6.9).

Table 6.7 shows the results of registering to the fluoroscopy image with soft tissue structures overlaid. The soft tissue structures caused a large failure rate when registering using the type '2' measures, all of which failed more than 50% of the time. Also, the valid registrations for the type '2' measures had large errors, $\overline{d}$ above 4.90mm. The failure rate for the type '1' and '3' measures also increased, but not by the same amount as the type '2' measures. The failure rate for local correlation and gradient correlation increased to 11.3% and 6.6% respectively, while the failure rate for pattern intensity and gradient difference increased hardly at all, to just 0.4%. The

Similarity measure	Type	RMSE rotation (degrees) θ_x	θ_y	θ_z	RMSE translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d_{(ROI)}}$ (mm)	No. fail, (%)
Cross correlation	2 B	1.15	2.14	1.49	21.2	1.77	1.75	4.90	1.56	50.4
Entropy	2 A	0.87	2.14	1.35	31.9	2.61	1.92	5.59	1.86	66.8
Mutual Information	2 A	1.55	2.25	1.66	22.1	2.30	0.99	5.62	1.93	77.3
Gradient correlation	3 B	0.30	0.56	0.47	3.9	0.48	0.35	1.36	0.58	6.6
Pattern Intensity	1 A	0.32	0.53	0.55	5.6	0.60	0.33	1.45	0.61	0.4
Gradient difference	3 A	0.30	0.62	0.50	4.9	0.42	0.30	1.41	0.60	0.4
Local correlation	1 A	0.47	0.56	0.60	5.9	0.47	0.39	1.63	0.64	11.3

Table 6.7: The performance of the similarity measures based on registrations to the fluoroscopy image of spine phantom with soft tissue structures overlaid.

registration errors for the type '1' and '3' measures increased slightly, with $\overline{d}$ increasing by about 0.3mm and $\overline{d_{(ROI)}}$ increasing by about 0.07mm.

Similarity measure	Type	RMSE rotation (degrees) θ_x	θ_y	θ_z	RMSE translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d_{(ROI)}}$ (mm)	No. fail, (%)
Cross correlation	2 B	1.55	2.44	1.91	24.6	1.54	1.58	6.25	1.79	42.2
Entropy	2 A	0.46	0.85	0.69	10.4	0.87	0.86	2.30	0.85	2.0
Mutual Information	2 A	0.83	1.58	1.28	22.8	1.25	1.13	3.89	1.24	37.1
Gradient correlation	3 B	0.56	0.63	0.42	8.3	0.34	0.25	1.59	0.60	14.5
Pattern Intensity	1 A	0.28	0.47	0.33	4.7	0.53	0.32	1.20	0.60	0.0
Gradient difference	3 A	0.24	0.45	0.35	4.9	0.44	0.36	1.13	0.55	0.0
Local correlation	1 A	0.44	0.58	0.48	6.4	0.57	0.47	1.60	0.66	12.1

Table 6.8: The performance of the similarity measures based on registrations to the fluoroscopy image of spine phantom with a stent overlaid.

When registrations were carried out to the fluoroscopy image with a stent overlaid (table 6.8) two of the similarity measures — cross correlation and mutual information — failed to register a number of times and the allowed registrations showed large errors. The results when using the gradient correlation and local correlation measures both show an increased failure rate and slightly larger errors. The results using the entropy, pattern intensity and gradient difference measures were largely unaffected by the presence of the interventional stent.

When registering to the image which contained both soft tissue structures and a stent (table 6.9), cross correlation, entropy and mutual information all failed a large number of times and final registrations showed large errors. Gradient correlation and local correlation failed 30% of the time, though final registrations were fairly accurate, $\overline{d_{(ROI)}}$ below 0.82mm. The gradient difference and pattern intensity measures were the most robust, with failure rates of 10.2% and 5.9% respectively,

Similarity measure	Type	RMSE rotation (degrees) θ_x	θ_y	θ_z	RMSE translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d(ROI)}$ (mm)	No. fail, (%)
Cross correlation	2 B	1.87	2.86	1.55	23.6	2.08	1.82	6.73	2.21	52.7
Entropy	2 A	0.77	2.15	1.75	32.0	2.40	1.95	5.86	2.07	78.1
Mutual Information	2 A	1.82	2.86	3.73	31.6	3.26	1.67	8.93	2.86	88.3
Gradient correlation	3 B	0.85	0.81	0.72	7.6	0.41	0.44	2.05	0.71	29.3
Pattern Intensity	1 A	0.50	0.48	0.51	6.8	0.68	0.31	1.52	0.63	5.9
Gradient difference	3 A	0.38	0.53	0.40	4.8	0.46	0.26	1.32	0.59	10.2
Local correlation	1 A	0.68	0.96	0.47	8.6	0.50	0.48	2.09	0.82	29.7

Table 6.9: How the similarity measures performed in registrations to fluoroscopy image of spine phantom with soft tissue and stent overlaid.

and the final registrations were the most accurate, $\overline{d} \leq 1.52$mm and $\overline{d(ROI)} \leq 0.63$mm.

6.3.2 Example values of the similarity measures

Example final values of the different similarity measures for successful and failed registrations are shown in table 6.10. These values were obtained for registrations to the L3 vertebra in the phantom fluoroscopy image with soft tissue structures and an interventional stent overlaid. All the values shown in table 6.10 should be maximised at registration. This was achieved by inverting the intensity values of the pixels in the fluoroscopy image ($I_{fl}(u,v) = I_{fl}(max) - I_{fl}(u,v)$) before calculating the similarity measures. The value for entropy of the difference image is the value of entropy in the fluoroscopy image before any features from the DRR have been subtracted $H(0)$ minus the final value of entropy of the difference image $H(s)$. Similarly the values given for pattern intensity and gradient difference both show the difference between the values of these measures when the scaling factor s (or s_V and s_H for gradient difference) equals zero and the final value of the similarity measure.

It should be noted that gradient difference and pattern intensity similarity measures are not normalised with respect to the size of the ROI in the fluoroscopy image and so the final values of these measures would change by a large amount if a different size ROI was used.

The results for the type 2 measures show that, for these particular set of registrations, there is not a large distinction between the final values of the similarity measures which found the registration position and those which found a local minima. For the cross correlation and entropy measures the difference between the mean values of the similarity measure for successful and failed registrations was not significant ($p>0.05$). For all of the type 1 and 3 measures there is a significant difference ($p>0.05$) between the mean values of the similarity measure for successful

Similarity measure	Type	Mean value of similarity measure		Value shown
		successful registrations	failed registrations	
Cross correlation	2B	0.545	0.546	R
Entropy	2A	0.208	0.200	$H(0) - H(s)$
Mutual Information	2A	0.491	0.480	S
Gradient correlation	3B	0.584	0.416	GC
Pattern Intensity	1A	8993	2433	$P(s) - P(0)$
Gradient difference	3A	20675	5708	$G(s_V, s_H) - G(0,0)$
Local correlation	1A	0.611	0.545	L

Table 6.10: Mean final values for the similarity measures for successful and failed registrations to the L3 vertebra in the phantom fluoroscopy image which contained both soft tissue and an interventional stent. All of the values shown should be maximised at registration. See section 5.2 for more details on the notation used for each similarity measure.

and failed registrations. This is especially evident for the pattern intensity and gradient difference measures.

6.3.3 Discussion

The introduction of clinical features into the fluoroscopy image of the phantom had an effect on the performance of all the similarity measures, but the magnitude of the effect was very different. The cross correlation, entropy and mutual information measures were affected a large amount. Gradient correlation and local correlation a small amount and pattern intensity and gradient difference were only marginally affected.

From table 6.7 it can be seen that the type '2' measures all have large failure rates and large RMS errors when registering to the image which contained overlaid soft tissue structures. This is to be expected, as these measures have no special attributes which would enable them to filter out the effects of soft tissue structures. They assume that a bony pixel in one part of the fluoroscopy image will have a similar intensity to a bony pixel in another part of the image. This need not be the case if there are underlying intensity fluctuations caused by soft tissue structures or the heel effect. The type '1' and '3' measures on the other hand are still able to register accurately and robustly despite the presence of soft tissue structures. They filter out the effect of soft tissue by two different methods, though they both make use of the same fundamental difference between bony and soft tissue detail — that soft tissue detail is on average of a lower spatial frequency than bony detail in fluoroscopy images.

When an interventional instrument is present in the fluoroscopy image, the pixels which contain the instrument have greatly altered intensity values. These pixels contain little to no useful information to aid the registration. The ability of a measure to register when an interventional stent

is present in the fluoroscopy image depends on whether the measure can effectively ignore these greatly altered pixels. Some similarity measures achieve this by making the maximum contribution from each pixel the same, i.e. independent of the intensity of the pixels. The number of pixels which contain features from interventional instruments are typically a small percentage of the total number of pixels in the fluoroscopy image, therefore, by making the maximum contribution from each pixel the same, the positive effect from the pixels which do not contain the instrument can by far outweigh any negative effects from the pixels which do contain the instrument. Table 5.2 splits the similarity measures up into two types depending on whether the maximum contribution from a pair of pixels is independent (type 'A') or dependent (type 'B') of the pixel intensity.

The results using the fluoroscopy image with a stent overlaid, table 6.8, shows that both of the type 'B' measures, cross correlation and gradient correlation, have an increased failure rate due to the presence of the stent, whereas, all of the type 'A' measures, except mutual information and local correlation, remain relatively unaffected by the stent. I believe that mutual information is failing again because the 2D histogram used by this measure is relatively sparsely populated. Local correlation, because it is a type 'A' measure, was expected to be robust to the presence of the stent. However, due to its regional nature and because the measure uses cross correlation, all of the pixels within a radius r of the stent would be rendered useless for registration purposes. This would greatly increase the number of useless pixels, which could be the cause of the higher failure rate when using the local correlation measure.

Table 6.9 shows the results when registering to the fluoroscopy image with both soft tissue structures and an interventional stent overlaid. The size of the errors and the number of failures increased for all the measures, but gradient difference and pattern intensity still remained accurate ($\overline{d(ROI)}$ equals 0.59mm and 0.63mm respectively) and reasonably robust (failure rate, 10.2% and 5.9% respectively).

Overall, the two most accurate similarity measures were pattern intensity and gradient difference. They were both able to achieve accurate registrations to a fluoroscopy image which contained both soft tissue structures and an interventional stent. The major difference between the measures is that pattern intensity reduces the effect of soft tissue by defining a region r within which there is little variation in intensity caused by soft tissue structures but a large variation in intensity caused by bony structures. The measure registers on the intensity values in a difference image, calculated by subtracting a scaled DRR from the fluoroscopy image. Gradient difference reduces the effect of soft tissue by applying Sobel templates to calculate gradient images, relying on there being more strong gradients due to bony features than due to soft tissue structures and air.

Registration is carried out on two gradient difference images, calculated by subtracting the scaled gradients in the DRR from the gradients in the fluoroscopy image.

6.4 Effect of slice thickness on registration accuracy.

The CT volume used in the previous experiments had a slice thickness of 1.5mm. This is a narrow slice thickness compared to standard clinical images which would typically have thicknesses ranging between 3mm to 10mm for abdominal imaging. The following experiment was carried out to ascertain the effect of the CT slice thickness on registration accuracy and robustness.

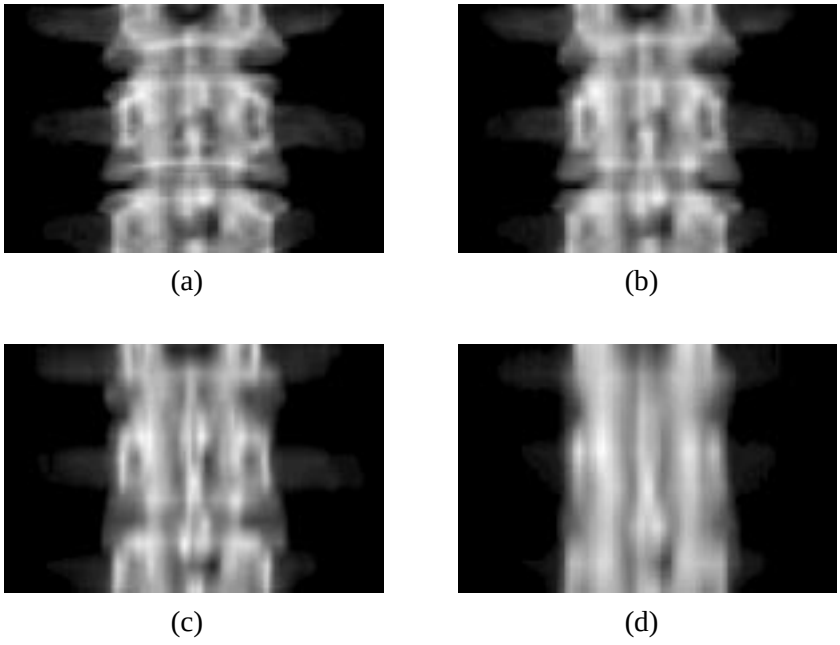


Figure 6.13: DRRs of spine phantom produced from CT volumes with different slice thicknesses (a) 1.5mm, (b) 3mm, (c) 6mm and (d) 10.5mm. DRR size 256x256 pixels.

6.4.1 Method

Registrations were all carried out to the L3 vertebra. The reduced CT volume around the L3 vertebra was used and its slice thickness was increased by averaging the voxel intensities between adjacent slices; the in-slice resolution not altered. Three additional reduced CT volumes were produced with slice thicknesses 3mm, 6mm and 10.5mm. Figure 6.13 shows the result of producing DRRs with these three CT volumes and also with the original reduced CT volume with a slice thickness of 1.5mm. The plain fluoroscopy image was also blurred with ellipses of minor and major axes 3x9 pixels, 3x19 pixels and 3x33 pixels to represent the new voxel resolutions. The

major axes of the ellipses were aligned with the vertical axis of the fluoroscopy image.

The same starting positions were used as before and sixty four registrations were carried out using each similarity measure and using each slice thickness.

6.4.2 Results

Table 6.11 shows **RMSE**, $\overline{d}$, $\overline{d(ROI)}$ and the failure rate when registering using different slice thickness and different similarity measures.

Similarity measure	Slice Thickness	**RMSE** rotation (degrees) θ_x	θ_y	θ_z	translation (mm) X	Y	Z	$\overline{d}$ (mm)	$\overline{d(ROI)}$ (mm)	No. fail, (%)
Cross Correlation	1.5 mm	0.62	0.78	0.48	8.7	0.54	0.90	2.09	1.08	0.0
	3 mm	0.81	0.44	0.41	9.1	0.41	1.16	2.20	1.25	0.0
	6mm	0.91	2.80	0.45	16.5	0.31	2.10	5.92	2.22	14.1
	10.5 mm	0.66	2.62	0.36	13.7	0.38	1.63	5.12	1.90	0.0
Entropy	1.5 mm	0.33	0.49	0.51	8.1	0.69	0.55	1.74	0.87	7.8
	3 mm	0.46	0.87	0.49	5.2	0.66	0.61	2.23	0.96	1.6
	6mm	0.43	3.20	0.69	8.7	0.75	0.76	5.79	1.69	28.1
	10.5 mm	0.35	2.57	0.49	11.8	0.46	1.02	4.81	1.49	0.0
Mutual Information	1.5 mm	0.68	1.40	0.58	24.1	0.56	1.29	3.31	1.30	17.2
	3 mm	0.77	1.58	0.51	23.9	0.57	1.32	4.06	1.49	21.9
	6mm	1.33	4.38	0.85	23.8	0.59	1.65	8.18	2.47	48.4
	10.5 mm	0.95	4.39	1.23	24.1	0.49	2.87	7.54	2.71	43.8
Gradient Correlation	1.5 mm	0.28	0.33	0.24	6.4	0.40	0.40	1.16	0.62	0.0
	3 mm	0.46	0.34	0.24	5.1	0.43	0.31	1.25	0.60	0.0
	6mm	0.46	1.56	0.22	10.2	0.29	0.22	2.98	0.85	0.0
	10.5 mm	0.41	4.56	0.25	13.3	0.29	0.26	7.68	2.00	1.6
Pattern Intensity	1.5 mm	0.29	0.25	0.33	4.9	0.44	0.38	1.16	0.63	0.0
	3 mm	0.45	0.38	0.31	5.3	0.43	0.29	1.34	0.60	0.0
	6mm	0.52	1.66	0.25	7.5	0.33	0.20	3.07	0.88	0.0
	10.5 mm	0.46	3.47	0.26	9.5	0.30	0.27	5.91	1.56	0.0
Gradient Difference	1.5 mm	0.29	0.25	0.29	6.0	0.38	0.40	1.13	0.61	0.0
	3 mm	0.36	0.49	0.25	4.8	0.41	0.27	1.22	0.57	0.0
	6mm	0.61	1.77	0.23	9.6	0.33	0.17	3.34	0.94	0.0
	10.5 mm	0.75	2.84	0.38	10.2	0.39	0.96	4.79	1.51	7.8
Local Correlation	1.5 mm	0.30	0.23	0.47	3.0	0.52	0.38	1.25	0.68	0.0
	3 mm	0.40	0.29	0.43	3.5	0.52	0.35	1.33	0.67	0.0
	6mm	0.51	1.37	0.34	7.7	0.41	0.33	2.75	0.86	0.0
	10.5 mm	0.60	3.47	0.33	8.3	0.37	0.31	5.98	1.59	0.0

Table 6.11: The effect on the accuracy and robustness of the similarity measures of altering the CT slice thickness.

6.4.3 Discussion

Table 6.11 shows the effect on accuracy and robustness of the similarity measures caused by altering the CT slice thickness. The first thing to note is that, as the slice thickness is increased, the largest changes in registration error occur in the θ_y and Z parameters. This is to be expected as the major effect of a Z translation or θ_y rotation is to cause structures in the DRR to move in the vertical direction. This direction is approximately perpendicular to the CT slice plane and therefore, this is the direction where information is lost from the CT volume as a result of increasing the CT slice thickness.

For all of the similarity measures the accuracy and robustness remained fairly constant as the slice thickness was increased from 1.5mm to 3mm. The accuracy, as measured by $\overline{d}$, roughly doubles when the slice thickness is increased to 6mm. Then, for similarity measures of type '1' or '3' (see table 5.1), the registration error increases further as the slice thickness is increased to 10.5mm. It is interesting that cross correlation and entropy both become more robust, and slightly more accurate, when the slice thickness is increased from 6mm to 10.5mm. The position of these failures when using the CT scan with a 6mm slice thickness and the cross correlation and entropy measures was investigated. The entropy measure found a single local minimum which was approximately a 14° rotation about θ_y away from the registration position. The cross correlation measure found two local minima which were approximately a 6° and a 15° rotation about θ_y away from the registration position. The increase in robustness for these two measures, when the slice thickness is increased from 6mm to 10.5mm, is believed to be because these local minima have been smoothed out at the 10.5mm slice thickness. As mentioned earlier the increase in slice thickness will smooth structures in the DRR in the vertical direction and the effect of a θ_y rotation is essentially to cause structures in the DRR to move in a vertical direction.

The registration accuracy, as measured by $\overline{d(ROI)}$, was below 1mm for all the similarity measures of type '1' or '3' when registering using the CT volume of slice thickness 1.5mm, 3mm and 6mm. However, $\overline{d(ROI)}$ increased to greater than 1.5mm for all the type '1' or '3' measures when using the CT scan with a 10.5mm slice thickness. Relating these results to my clinical requirement (for the aortic aneurysm or percutaneous laser nucleotomy procedures) of 1mm accuracy indicates that the CT slice thickness should not be greater than 6mm.

6.5 Effect of altering the out-of-plane translation direction.

As described in section 5.4.3 the out-of-plane translation direction for the algorithm is along a line from the x-ray source to the centre of the CT volume (direction B in figure 5.8). The experiment described in this section was carried out to demonstrate the advantage of translating along a line from the x-ray source to the centre of the CT volume, rather than along a direction parallel to the X-axis (direction A in figure 5.8). An experiment was carried out to register each of the five lumbar vertebrae to the plain fluoroscopy image using the pattern intensity similarity measure. Firstly the standard algorithm was used, where the out-of-plane translation direction is along a line from the x-ray source and the centre of the CT volume. Secondly, the algorithm was modified so that the out-of-plane translation direction was parallel to the X-axis and the experiment was repeated. The pattern intensity measure was chosen because it was one of the most accurate and robust in previous experiments.

6.5.1 Results

Results are presented in the following two tables. Table 6.12 shows $\mathbf{RMSE}$, $\overline{d}$, $\overline{d(ROI)}$ and the failure rate when the out-of-plane translation direction is along an axis defined by a vector from the source to the centre of the CT volume. Table 6.13 shows $\mathbf{RMSE}$, $\overline{d}$, $\overline{d(ROI)}$ and the failure rate when the out-of-plane translation direction is parallel to the X-axis.

Vertebra	$\mathbf{RMSE}$						$\overline{d}$	$\overline{d(ROI)}$	No.
	rotation (degrees)			translation (mm)			(mm)	(mm)	fail,
	θ_x	θ_y	θ_z	X	Y	Z			(%)
L1	0.42	0.19	0.29	8.5	0.53	0.77	1.48	0.41	0.0
L2	0.25	0.22	0.37	6.0	0.61	0.54	1.21	0.64	0.0
L3	0.17	0.27	0.34	5.2	0.64	0.40	1.19	0.65	0.0
L4	0.19	0.49	0.21	4.7	0.39	0.20	1.04	0.57	0.0
L5	0.17	0.36	0.30	2.5	0.38	0.30	0.91	0.43	0.0

Table 6.12: Results from pattern intensity similarity measure when the out-of-plane translation direction is along a vector from the x-ray source to the centre of the CT volume.

6.5.2 Discussion

The registration results when the out-of-plane translation direction for the algorithm is parallel to the X-axis, see table 6.13, were all less accurate (as defined by $\overline{d}$ and $\overline{d(ROI)}$) than the registration results when the out-of-plane translation was along a line from the x-ray source to the centre of the CT volume. The X-axis intercepts the fluoroscopy screen inside the projected image of the L3

Vertebra	RMSE						$\overline{d}$ (mm)	$\overline{d(ROI)}$ (mm)	No. fail, (%)
	rotation (degrees)			translation (mm)					
	θ_x	θ_y	θ_z	X	Y	Z			
L1	0.59	0.35	0.34	32.4	1.18	2.82	4.15	0.74	3.1
L2	0.18	0.32	0.24	19.3	0.62	0.86	2.01	0.68	1.6
L3	0.33	0.35	0.41	10.2	0.81	0.42	1.57	0.68	1.6
L4	0.23	0.38	0.46	17.9	0.74	0.71	1.76	0.66	0.0
L5	0.17	0.69	0.45	33.8	1.12	2.06	3.98	0.88	9.4

Table 6.13: Results from pattern intensity similarity measure when the out-of-plane translation direction is parallel to the X-axis

vertebra. From table 6.13 the registration error, particularly in the X direction can be seen to be at a minimum for the L3 vertebra, then increase for the neighbouring vertebrae (L2 and L4) and then increase further for registrations to L1 and L5. This is to be expected because, as the ROI moves away from the position where the X-axis intercepts with the fluoroscopy image, more local minima will occur because the translational parameters are not decoupled, see figure 5.7.

6.6 Effect of allowing non-physical rays

In the production of DRRs non-physical rays are detected and are not included in the DRR. These non-physical rays can arise due to the truncated nature of the CT volume, where rays enter or exit through the top or base of the CT volume, see section 4.3.2. These rays have no physical counterpart in a fluoroscopy image and so should not be compared with such an image. Non-physical rays can also arise from another situation, when structures are segmented from the CT volume. For example figure 6.14(b) shows a DRR produced from a CT scan which has been segmented to remove everything except the L3 vertebra. In this case non-physical rays occur if a ray which intercepts the segmented structure, would have also intercepted another bony structure had the CT volume not been segmented. The non physical rays shown in figure 6.14(b) include the rays which go through the inferior articular processes and the spinous process from the L3 vertebra, these rays should have also travelled through the L4 vertebral body.

To investigate the effect of non-physical rays the L3 vertebra was manually segmented from the CT volume. Registrations were carried out using the segmented CT volume with no ROI defined in the fluoroscopy image. Sixty four registrations were carried using each similarity measure with the same starting positions as before.

6.6.1 Results

Results are presented in the following tables. They show **RMSE**, $\overline{d}$ and the failure rate for each similarity measure. The $\overline{d_{(ROI)}}$ error measurement was not calculated for these results because the registrations did not use a ROI within the fluoroscopy image. Firstly, for comparison, table 6.14 shows the registration results when the standard L3 reduced CT volume and ROI were used. Then table 6.15 shows the results of registering on the segmented L3 vertebra with no ROI in the fluoroscopy image.

Similarity measure	Type	**RMSE** rotation (degrees)			**RMSE** translation (mm)			$\overline{d}$ (mm)	No. fail, (%)
		θ_x	θ_y	θ_z	X	Y	Z		
Cross correlation	2 B	0.62	0.78	0.47	8.7	0.54	0.90	2.09	0.0
Entropy	2 A	0.33	0.49	0.51	8.1	0.69	0.55	1.74	7.8
Mutual Information	2 A	0.68	1.40	0.58	24.1	0.56	1.29	3.31	17.2
Gradient correlation	3 B	0.28	0.33	0.24	6.4	0.40	0.40	1.16	0.0
Pattern Intensity	1 A	0.29	0.25	0.33	4.9	0.44	0.38	1.16	0.0
Gradient difference	3 A	0.29	0.33	0.24	6.4	0.40	0.40	1.13	0.0
Local correlation	1 A	0.30	0.23	0.47	3.0	0.52	0.38	1.25	0.0

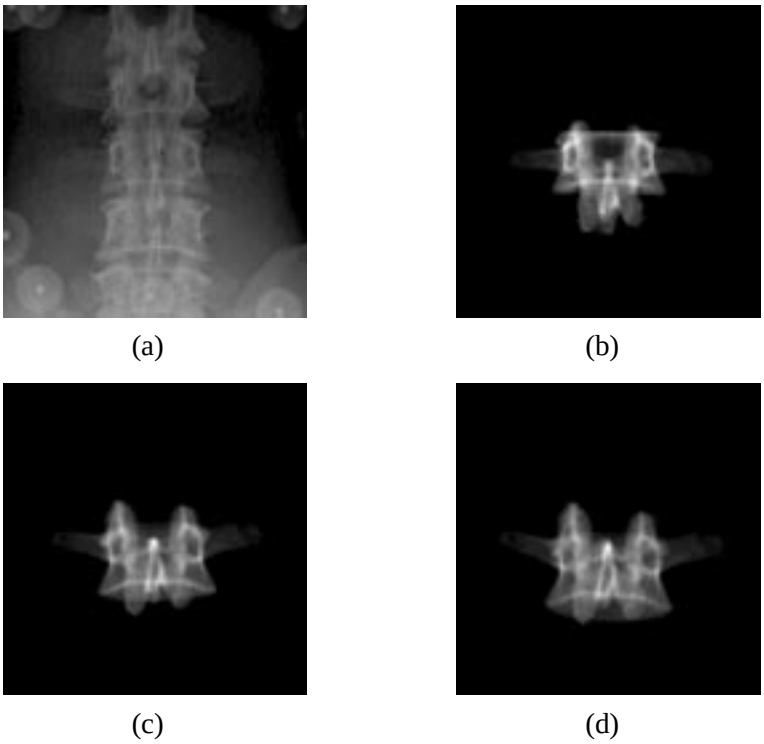
Table 6.14: The performance of the similarity measures based on registrations between the reduced CT volume and the fluoroscopy image with a ROI around the L3 vertebra.

Similarity measure	Type	**RMSE** rotation (degrees)			**RMSE** translation (mm)			$\overline{d}$ (mm)	No. fail, (%)
		θ_x	θ_y	θ_z	X	Y	Z		
Cross correlation	2 B	2.02	3.55	1.28	25.4	0.95	0.33	7.00	75.0
Entropy	2 A	1.92	1.71	1.04	27.7	0.36	1.55	4.94	73.4
Mutual Information	2 A	0.84	1.57	2.07	28.4	1.04	1.07	5.14	81.3
Gradient correlation	3 B	0.40	0.41	0.19	4.2	0.43	0.25	1.13	0.0
Pattern Intensity	1 A	0.27	0.45	0.18	7.5	0.39	0.44	1.23	0.0
Gradient difference	3 A	0.28	0.43	0.19	5.7	0.38	0.27	1.12	0.0
Local correlation	1 A	0.37	0.77	0.70	7.4	0.51	0.39	2.12	1.6

Table 6.15: The performance of the similarity measures based on registrations between a segmented CT volume which just contains the L3 vertebra and the fluoroscopy image with no ROI specified.

6.6.2 Discussion

The results for the type '1' and type '3' measures (apart from local correlation) showed very little change in registration accuracy. The effect of including non-physical rays caused the $\overline{d}$ error measure to increase on average by about 0.1mm.



(a) (b)

(c) (d)

Figure 6.14: Fluoroscopy image of spine phantom (a) and DRRs at final registration positions using pattern intensity (b), cross correlation (c) and entropy (d).

This investigation produced some surprising results for the type '2' measures. Allowing non-physical rays resulted in a large increase in the failure rate and in $\overline{d}$ for these similarity measures. Figure 6.14(c) and (d) show typical final registration positions when using the cross correlation and entropy similarity measures when non-physical rays are allowed. These two measures appear to be failing for different reasons. Firstly, the reasons why the cross correlation measure is failing are discussed. Cross correlation is calculated over the region of overlap. At the registration position, figure 6.14(b), the inferior articular processes and the spinous process from the L3 vertebra are overlying a number of features from the L4 vertebral body. Therefore, the cross correlation value between pixel intensities from the processes in the DRR and the corresponding areas in the fluoroscopy image will be low. By applying a rotation about θ_y (see figure 3.1) the algorithm is able to hide the processes behind the L3 vertebral body, see figure 6.14(c). The region with the low cross correlation value has been removed and the overall correlation value has increased, even though the DRR and fluoroscopy images are now misregistered. These effects may also be responsible for the decrease in registration accuracy of the local correlation measure.

Secondly, the reasons why the entropy measure is failing are discussed. Entropy attempts to remove structure from the difference image by overlaying corresponding structures from each im-

age. If there is no ROI in the fluoroscopy image then there are a much larger number of structures in the fluoroscopy image with no counterpart in the DRR. However, the structures in the DRR can change size during the registration procedure by translating towards the x-ray source. In the typical final registration position using the entropy measure (figure 6.14(d)) the vertebra has been rotated to hide the processes behind the vertebral body and it has also been translated towards the source, thereby increasing the size of the structures in the DRR and so removing more structure from the difference image. In figure 6.14(d) the DRR is larger than at "gold-standard' registration position and it is also overlying part of the L4 vertebra. The value of entropy of the difference image calculated using the DRR in figure 6.14(d) is lower than the entropy value calculated using a DRR produced at the "gold-standard" registration position, figure 6.14(b).

6.7 Conclusions

This chapter has compared the performance of seven similarity measures when registering a CT scan to a fluoroscopy image of a spine phantom. The final registration positions from the algorithm were compared in each case to a "gold-standard" registration calculated using fiducial markers. This "gold-standard" registration has been calculated to be accurate to a RMS error of 1.2 pixels at the imaging plane, which has been shown to correspond to a standard deviation of 0.09° in the rotational parameters, 0.15mm for the in-plane translational parameters and 1.34mm for the out-of-plane translational parameter.

This is the first time such an extensive comparison of intensity based similarity measures for 2D-3D registration has been carried out. It is also the first time that the effect on the performance of different similarity measures, of features such as soft tissue structures and interventional instruments has been studied. The results in this chapter have shown that these features can have a large effect on the accuracy and robustness of some similarity measures which have been previously used for 2D-3D image registration.

All of the measures, except for mutual information registered accurately and robustly to the plain fluoroscopy image. Similarity measures which do not use a local region or gradient images were shown to fail when soft tissue structures were introduced into the phantom fluoroscopy image. Similarity measures which use cross correlation were shown to fail when an interventional stent was introduced into the phantom fluoroscopy image. The most accurate and robust similarity measures were pattern intensity and gradient difference. These measures had an overall failure rate for registrations to the plain and the simulated fluoroscopy images of only 1.6% and 2.7%, and an

average registration accuracy, as measured by $\overline{d(ROI)}$, of only 0.60mm and 0.58mm respectively.

This section has also investigated the effect of increasing slice thickness on registration accuracy. The results indicated that, to achieve a registration accuracy of less than 1mm, the slice thickness should not be greater than 6mm. The importance of decoupling the in-plane and out-of-plane translations was shown in section 6.5: and in section 6.6 the inclusion of non-physical rays was shown to have a large effect on type '2' measures, but have a negligible effect on the other measures.

Chapter 7

An experimental comparison of similarity measures using clinical images

7.1 Introduction

The previous chapter has investigated the accuracy and robustness of seven similarity measures in registering a CT volume to a fluoroscopy image of a spine phantom. Although the images used were of a spine phantom, clinical images were simulated by overlaying features segmented from clinical fluoroscopy images. This chapter aims to extend the previous investigation by investigating the accuracy and robustness of the similarity measures using routinely acquired clinical images from aortic stenting procedures. A major problem when using clinical images is that it is extremely difficult to obtain an accurate "gold-standard" registration. It is relatively easy to attach fiducial markers to a rigid phantom, however, the attachment of fiducial markers to patients, unless it is already part of the clinical protocol, is fraught with difficulties. The use of fiducial markers (or stereotactic frames) for procedures carried out using fluoroscopy guidance is extremely rare. Exceptions are in stereotactic neurosurgery [64] and the ROBODOC$^{(R)}$ system [13, 119] for hip replacement surgery. Clinical images of the lumbar spine which included fiducial markers were not available for the work described in this chapter and so images from aortic stenting procedures were used. Because of this it was not possible to calculate a "gold-standard" registration. Instead the registration accuracy has been estimated using a *k-fold* cross validation [66] approach.

The first section in this chapter explains how *k-fold* cross validation can be used as a measure

of registration accuracy. Registrations are carried out to each of the vertebrae present in a given set of fluoroscopy and CT images. One vertebra is chosen and the registrations positions to this vertebra are compared to a mean set of registration positions calculated using the other vertebrae. This process is repeated and a different vertebra is chosen each time. However, this method will be affected if spinal deformation has occurred between when the CT image and when the fluoroscopy image were acquired. To account for spinal deformation two fluoroscopy views, taken from different positions, are used. Individual registrations are carried out to each vertebra in each fluoroscopy image. The registration results to the same vertebra, but in each fluoroscopy image, are used to compute the transformation of the fluoroscopy set between the acquisition of the two fluoroscopy images. Assuming that no spinal deformation has occurred between the acquisition of the two fluoroscopy images (and the registrations are accurate) then the transformation of the fluoroscopy set calculated using different vertebrae should be identical. The *k-fold* cross validation approach is used to calculate how consistently the registration algorithm is able to calculate the transformation of the fluoroscopy set using registrations to different vertebrae. This is shown in the subsequent sections to be a reasonable approximation of the registration accuracy of the algorithm.

An experiment is carried out using the spine phantom dataset with a known "gold-standard" registration to show that the *k-fold* cross validation method can be used to estimate RMS registration error. The next section contains the main experimental part of this chapter, which is an investigation into the accuracy and robustness of seven similarity measures using four routinely acquired clinical datasets. The final section in this chapter carries out registrations using images from a percutaneous laser nucleotomy procedure. Pre-operative CT and MR images were used to plan this procedure. By registering the pre-operative MR image to the pre-operative CT image using normalised mutual information [112, 116] it was possible to overlay information from the pre-operative MR image onto the intra-operative fluoroscopy image.

7.2 Using *k-fold* cross validation to estimate registration error

This section begins with an overview on how *k-fold* cross validation can be used to estimate RMS registration errors. By carefully designing the experiment a number of independent datasets have been generated which should all have the same mean values. Because these datasets are independent the error of the mean value for each dataset should also be independent. This section then explains how the effects of spinal deformation are accounted for by using two fluoroscopy images. The final part of this section discusses the expected causes of error in my method.

7.2.1 Overview: estimating RMS errors using *k-fold* cross validation

In this chapter a *k-fold* cross validation [66] approach has been used to estimate registration errors. This approach involves using the results from N_{good} registrations to obtain a dataset $\mathcal{D} = \{\mathbf{E}_1, \ldots, \mathbf{E}_{N_{good}}\}$. $\mathcal{D}$ is split up into a training ($\mathcal{D}_{train}$) and a testing dataset ($\mathcal{D}_{test} = \mathcal{D} \setminus \mathcal{D}_{train}$). $\mathcal{D}_{train}$ is used to calculate a mean set of extrinsic parameters $\overline{\mathbf{E}}$ as shown in equation (7.1).

$$\overline{E}^i = \frac{1}{N_{train}} \sum_{j=1}^{N_{train}} E_j^i \ , \quad i = 1, \ldots, 6 \tag{7.1}$$

Where the superscript i represents each of the individual extrinsic parameters in $\mathbf{E}$ and the subscript j represents the elements of the training dataset $\mathcal{D}_{train}$ which is of size N_{train}.

The root mean square (RMS) difference $\mathbf{RMSD}$ between $\overline{\mathbf{E}}$ and the elements of the set $\mathcal{D}_{test}$ is calculated using equation (7.2).

$$RMSD^i = \sqrt{\frac{1}{N_{test}} \sum_{j=1}^{N_{test}} (E_j^i - \overline{E}^i)^2} \ , \quad i = 1, \ldots, 6 \tag{7.2}$$

Where E_j now represents the elements of the test dataset $\mathcal{D}_{test}$ which is of size $N_{test} = N_{good} - N_{train}$.

The dataset $\mathcal{D}$ is split up n times into $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$, and the overall error measure calculated using *k-fold* cross validation, $\mathbf{CV}$, is given by equation (7.3).

$$CV^i = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (RMSD_j^i)^2} \ , \quad i = 1, \ldots, 6 \tag{7.3}$$

Where $\mathbf{RMSD}_j$ represents the RMS difference calculated using the j'th split of the dataset.

7.2.2 Method assuming no spinal deformation.

A problem can occur when using cross validation to estimate RMS errors if there is a similar, non zero, bias in the testing and training datasets. The term bias is used here to represent the signed difference between the mean extrinsic parameters in the datasets and a set of "true" or "gold-standard" extrinsic registration parameters. If such a bias occurs then equation (7.3) approximately calculates the standard deviation of the dataset $\mathcal{D}$. If such a bias does not occur then the cross validation error value $\mathbf{CV}$ should be a reasonable approximation to a RMS error value measured using an independent "gold-standard" registration. To counter the possible problems which could be caused by a constant bias, care has been taken in the design of the experiment so that the test and training datasets are formed from registrations using different regions of the CT and

fluoroscopy images. The assumption is that, because different regions of the images are used the input information into the registration algorithm is different for each dataset and so the errors in the mean values of the testing and training datasets should also be independent.

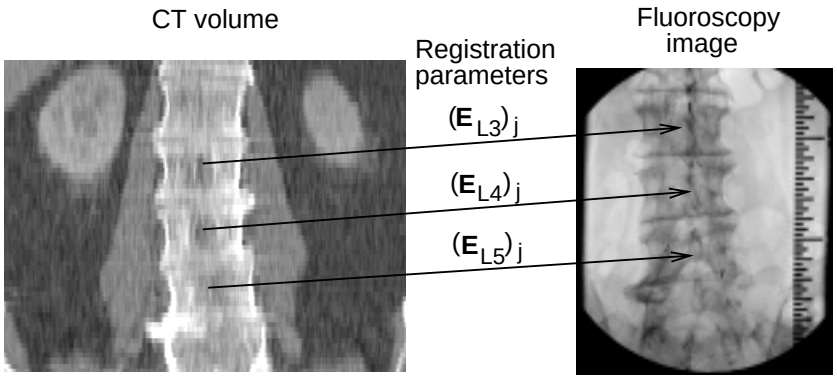


Figure 7.1: Cross validation method assuming no spinal deformation. Sixty four registrations are carried out using each vertebra to give three sets of final extrinsic parameters $(\mathbf{E}_{L3})_j$, $(\mathbf{E}_{L4})_j$ and $(\mathbf{E}_{L5})_j, j = 1, \ldots, 64$. One set of parameters (e.g. $\mathbf{E}_{L3}$) is then used as the test data set and the other two sets of parameters (e.g. $\mathbf{E}_{L4}$ and $\mathbf{E}_{L5}$) form the training data set.

The method, assuming no spinal deformation, is shown schematically in figure 7.1. A clinical CT volume, represented by a coronal CT slice in figure 7.1, and a fluoroscopy image of the lumbar spine are used. A ROI is defined for each of the lumbar vertebra in the fluoroscopy image and a reduced CT volume is defined around each vertebra in the CT volume. Sixty four registrations are carried out to each vertebra and the final registration parameters are represented by $(\mathbf{E}_{L3})_j$, $(\mathbf{E}_{L4})_j$ and $(\mathbf{E}_{L5})_j, j = 1, \ldots, 64$. These parameters are firstly filtered to remove failures and are then transformed from a coordinate system based on a reduced CT volume to a coordinate system based on the whole CT volume. If there is no spinal deformation and the registrations are accurate, then the transformed final extrinsic parameters from each vertebra should be identical. These registration results are input into equations (7.1), (7.2) and (7.3) to estimate the RMS error using cross validation, $\mathbf{CV}$. The final extrinsic parameters are split up by removing the results from registrations to one vertebra as $\mathcal{D}_{test}$, e.g. removing $\mathbf{E}_{L3}$. The number of successful registrations to this vertebra determines the size of the test dataset, N_{test}. The results from the other vertebrae form the training dataset, $\mathcal{D}_{train}$, e.g. $\mathbf{E}_{L4}$ and $\mathbf{E}_{L5}$. The results from registrations to each vertebrae are removed in turn, therefore, in figure 7.1 where three lumbar vertebrae are used, $n = 3$.

7.2.3 Accounting for spinal deformation.

If spinal deformation has occurred between the acquisition of the CT image and the fluoroscopy image, then the final extrinsic parameters from registering using each vertebra may be different and so the above method to estimate registration accuracy cannot be used. Instead the following method, which is not affected by spinal deformation, is used. This method accounts for spinal deformation by using two fluoroscopy images taken from different positions (e.g. AP and lateral), see figure 7.2. It is assumed that no spinal deformation has occurred between the two fluoroscopy images. This is a reasonable assumption as the patient is not moved between the fluoroscopy image acquisitions. The aim is to find a transformation, based upon the registration results to individual vertebra, which will not be affected by spinal deformation between the CT and the fluoroscopy images. The transformation of the fluoroscopy set, $\mathbf{F}$, between when the two fluoroscopy images were acquired, is such a transformation. The *k-fold* cross validation approach is applied to the parameters defining the transformation of the fluoroscopy set. The method used is as follows.

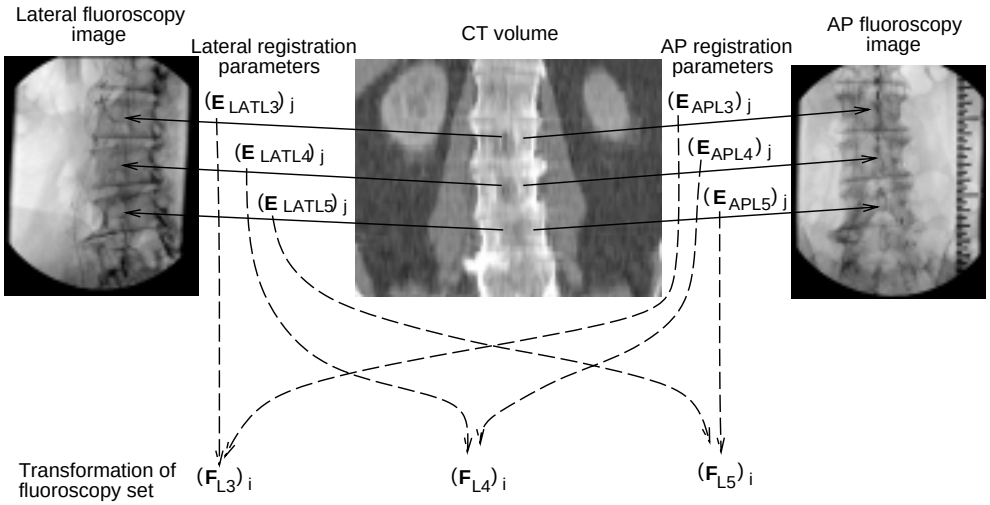


Figure 7.2: Cross validation method which accounts for deformation between the CT and the fluoroscopy images. Sixty four registrations are carried out to each vertebrae in each fluoroscopy image. Note, only one fluoroscopy image is used for any individual registration. The final registration results to the same vertebra but in each fluoroscopy image (e.g. $\mathbf{E}_{APL3}$ and $\mathbf{E}_{LATL3}$) are used to calculate the transformation of the fluoroscopy set between the AP and the lateral view (e.g. $\mathbf{F}_{L3}$). The *k-fold* cross validation procedure is applied to the parameters describing the transformation of the fluoroscopy set based on registrations to different vertebrae, e.g. $\mathbf{F}_{L3}$, $\mathbf{F}_{L4}$ and $\mathbf{F}_{L5}$.

Individual registrations are carried out to each vertebra in each fluoroscopy image, therefore, as shown in figure 7.2 if three lumbar vertebrae are used we obtain six sets of final extrinsic pa-

rameters, $(\mathbf{E}_{APL3})_j$, $(\mathbf{E}_{APL4})_j$, $(\mathbf{E}_{APL5})_j$, $(\mathbf{E}_{LATL3})_j$, $(\mathbf{E}_{LATL4})_j$ and $(\mathbf{E}_{LATL5})_j, j = 1, \ldots, 64$. Please note that registrations are only ever carried out to a single fluoroscopy image, the two fluoroscopy views are never used at the same time by the registration algorithm. The final extrinsic parameters from registrations to the same vertebra, but using different fluoroscopy images, are used to compute the transformation of the fluoroscopy set, $\mathbf{F}$. For example, the sets of extrinsic parameters $\mathbf{E}_{APL3}$ and $\mathbf{E}_{LATL3}$ are used to calculate the set of parameters defining the movement of the fluoroscopy set based on registrations to the L3 vertebra. Section 7.3 describes how the parameters defining the transformation of the fluoroscopy set, $\mathbf{F} = (\theta_x, \theta_y, \theta_z, X, Y, Z)$, are calculated and shows that these parameters are independent of any spinal deformation. The *k-fold* cross validation procedure is applied to a dataset consisting of N different transformations of the fluoroscopy set i.e. $\mathcal{D} = \{\mathbf{F}_1, \ldots, \mathbf{F}_N\}$. The results from one vertebra are removed from the total set of transformations of the fluoroscopy set to create $\mathcal{D}_{test}$ and $\mathcal{D}_{train}$ is formed using the results from the remaining vertebrae. This is again repeated three times, removing the results from a different vertebra each time. Equations (7.1), (7.2) and (7.3) are used to obtain an estimate of the RMS error in the parameters defining the transformation of the fluoroscopy set. This RMS error is comprised from errors in both the AP and lateral registrations. The estimated error in a single registration, assuming the two registrations to be independent and identically distributed, will be $1/\sqrt{2}$ times smaller. The error measure reported in the following sections is $\mathbf{CV}'$, it is based on the estimated RMS error of the transformation of the fluoroscopy set multiplied by $1/\sqrt{2}$, i.e. $\mathbf{CV}' = \mathbf{CV}/\sqrt{2}$.

7.2.4 Expected causes of error in using a *k- fold* cross validation technique to estimate RMS errors

As mentioned previously, if there are similar, non zero errors in the mean values of the test and training datasets, then these will not be detected by the cross validation technique. There are two types of similar bias which would not be detected using my implementation of the cross validation technique. The first type occurs if all the registrations to different vertebrae in the same fluoroscopy image are biased in a similar way (e.g. the final mean registration position for all of the vertebrae in one image is always shifted too close to the x-ray source). The second type can occur because two fluoroscopy images are used. If registrations to the same vertebrae in the AP and the lateral view are biased in a similar way (e.g. the final mean registration position for the L3 vertebra in both the AP and lateral images are too close to the x-ray source).

In designing the experiment I have tried to reduce the possibility of a similar bias occurring

in the test and training datasets by using information from different vertebrae to create the test and training datasets. Therefore, the image information used to create the testing datasets is independent of the image information used to create the training datasets. However, there may be other sources of bias, possibly due to the large degree of similarity in the shape of different vertebrae. To investigate whether other sources of constant bias might be present, which could effect the results, an experiment has been carried out using phantom images which have a known "gold-standard". This experiment is described in section 7.4.

7.3 Calculating the parameters defining the transformation of the fluoroscopy set.

This section describes how the parameters defining the transformation of the fluoroscopy set, $\mathbf{F}$, between an AP and a lateral view can be calculated from the two sets of parameters $\mathbf{E}_{AP}$ and $\mathbf{E}_{LAT}$ which represent the final registration positions between a CT scan and the AP and lateral fluoroscopy images. The parameters $\mathbf{E}_{AP}$, $\mathbf{E}_{LAT}$ and $\mathbf{F}$ can be used to calculate the rotational ($\mathbf{R}_{AP}$, $\mathbf{R}_{LAT}$, $\mathbf{R}_{fl}$) and translational ($\mathbf{t}_{AP}$, $\mathbf{t}_{LAT}$, $\mathbf{t}_{fl}$) matrices which are used in the following calculations.

Figure 7.3 shows the three coordinate systems used in this section. The first coordinate system is based on the pre-operative CT volume where $\mathbf{x}$ represents a position in the CT volume. The other two coordinate systems are defined by the fluoroscopy gantry. Where $\mathbf{x}_{AP}$ represents a position relative to the fluoroscopy gantry when the AP image was acquired and $\mathbf{x}_{LAT}$ represents the position relative to the fluoroscopy gantry when the lateral image was acquired. The transformation matrices between these three coordinate systems are shown in figure 7.3. The aim of this section is to calculate, $\mathbf{F}$, which represents the parameters defining the movement of the fluoroscopy set between an AP and a lateral position. These parameters can be obtained by decomposing the rotation and translation matrices governing the movement of the fluoroscopy set, $\mathbf{R}_{fl}$ and $\mathbf{t}_{fl}$.

Assuming no spinal deformation, the transformation between the CT coordinate system and the AP or lateral coordinate systems are given by equations (7.4) and (7.5), where $\mathbf{R}$ and $\mathbf{t}$ represent rotational and translational matrices respectively.

$$\mathbf{x}_{AP} = \mathbf{R}_{AP}\mathbf{x} + \mathbf{t}_{AP} \tag{7.4}$$

$$\mathbf{x}_{LAT} = \mathbf{R}_{LAT}\mathbf{x} + \mathbf{t}_{LAT} \tag{7.5}$$

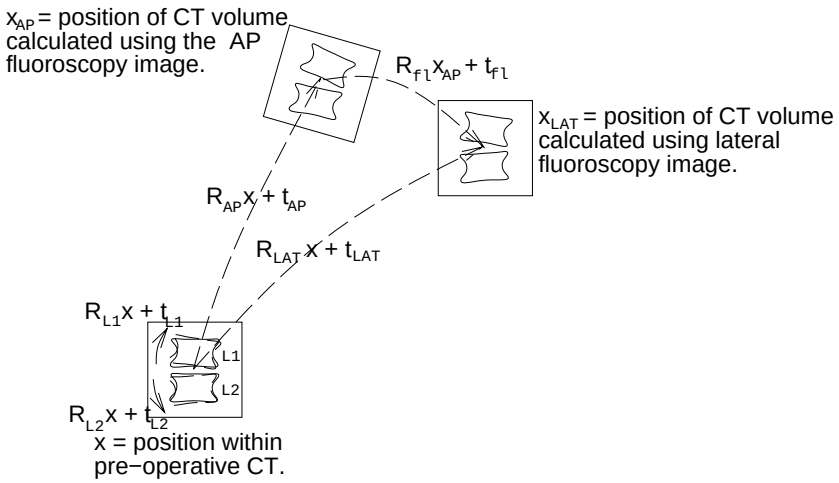


Figure 7.3: The coordinate systems and transformation matrices used to calculate the movement of the fluoroscopy set. There are three coordinate systems; $\mathbf{x}$ represents the position within the pre-operative CT image; $\mathbf{x}_{AP}$ and $\mathbf{x}_{LAT}$ represent the position, relative to the fluoroscopy gantry, when the gantry is in the AP and lateral positions respectively. The rotational and translational matrices between these coordinate systems are represented by $\mathbf{R}_{AP}, \mathbf{t}_{AP}, \mathbf{R}_{LAT}, \mathbf{t}_{LAT}, \mathbf{R}_{fl}$ and $\mathbf{t}_{fl}$. Spinal deformation is represented by the matrices $\mathbf{R}_{L1}, \mathbf{t}_{L1}, \mathbf{R}_{L2}$ and $\mathbf{t}_{L2}$. These matrices transform the L1 and L2 vertebrae from their pre-operative position to their intra-operative position.

Spinal deformation is also depicted in figure 7.3. The matrices $\mathbf{R}_{L1}, \mathbf{t}_{L1}, \mathbf{R}_{L2}$ and $\mathbf{t}_{L2}$ define the transformation of the L1 and L2 vertebrae from their pre-operative position to their intra-operative position. Combining these transformations, where the pre-operative volume is firstly deformed and then transformed into $\mathbf{x}_{AP}$ and $\mathbf{x}_{LAT}$, gives the following four equations, (7.6) to (7.9). These equations represent the transformations of each vertebra from the pre-operative CT to the two coordinate systems defined by the position of the fluoroscopy gantry as the AP and lateral images are acquired. The rotational and translational matrices $\mathbf{R}_{APL1}, \ldots, \mathbf{R}_{LATL2}$, $\mathbf{t}_{APL1}, \ldots, \mathbf{t}_{LATL2}$ represent the transformation including deformation (e.g. $\mathbf{R}_{APL1} = \mathbf{R}_{AP}\mathbf{R}_{L1}$ and $\mathbf{t}_{APL1} = \mathbf{R}_{AP}\mathbf{t}_{L1} + \mathbf{t}_{AP}$). These matrices can be calculated using the final extrinsic parameters from the registration algorithm, e.g. the final extrinsic parameters when registering to the L1 vertebra in the AP view ($\mathbf{E}_{APL1}$) can be used to calculate $\mathbf{R}_{APL1}$ and $\mathbf{t}_{APL1}$.

$$\mathbf{x}_{APL1} = \mathbf{R}_{AP}\mathbf{R}_{L1}\mathbf{x} + \mathbf{R}_{AP}\mathbf{t}_{L1} + \mathbf{t}_{AP} = \mathbf{R}_{APL1}\mathbf{x} + \mathbf{t}_{APL1} \tag{7.6}$$

$$\mathbf{x}_{APL2} = \mathbf{R}_{AP}\mathbf{R}_{L2}\mathbf{x} + \mathbf{R}_{AP}\mathbf{t}_{L2} + \mathbf{t}_{AP} = \mathbf{R}_{APL2}\mathbf{x} + \mathbf{t}_{APL2} \tag{7.7}$$

$$\mathbf{x}_{LATL1} = \mathbf{R}_{LAT}\mathbf{R}_{L1}\mathbf{x} + \mathbf{R}_{LAT}\mathbf{t}_{L1} + \mathbf{t}_{LAT} = \mathbf{R}_{LATL1}\mathbf{x} + \mathbf{t}_{LATL1} \tag{7.8}$$

$$\mathbf{x}_{LATL2} = \mathbf{R}_{LAT}\mathbf{R}_{L2}\mathbf{x} + \mathbf{R}_{LAT}\mathbf{t}_{L2} + \mathbf{t}_{LAT} = \mathbf{R}_{LATL2}\mathbf{x} + \mathbf{t}_{LATL2} \tag{7.9}$$

The transformation of the fluoroscopy set, $\mathbf{R}_{fl}$ and $\mathbf{t}_{fl}$, based on registrations to the L1 vertebra can be calculated by removing $\mathbf{x}$ from equations (7.6) and (7.8). This is shown in equation (7.10).

$$\mathbf{x}_{LATL1} = \mathbf{R}_{LATL1}\mathbf{R}^T_{APL1}\mathbf{x}_{APL1} + \mathbf{t}_{LATL1} - \mathbf{R}_{LATL1}\mathbf{R}^T_{APL1}\mathbf{t}_{APL1} = \mathbf{R}_{fl}\mathbf{x}_{APL1} + \mathbf{t}_{fl} \tag{7.10}$$

The rotation matrix from equation (7.10), $\mathbf{R}_{fl}$, is shown in equation (7.11) to be independent of spinal deformation.

$$\mathbf{R}_{fl} = \mathbf{R}_{LATL1}\mathbf{R}^T_{APL1} = \mathbf{R}_{LAT}\mathbf{R}_{L1}(\mathbf{R}_{AP}\mathbf{R}_{L1})^T = \mathbf{R}_{LAT}\mathbf{R}^T_{AP} \tag{7.11}$$

Similarly the translational matrix from equation (7.10), $\mathbf{t}_{fl}$, is shown by equations (7.12) to (7.14) to also be independent of spinal deformation.

$$\mathbf{t}_{fl} = \mathbf{t}_{LATL1} - \mathbf{R}_{LATL1}\mathbf{R}^T_{APL1}\mathbf{t}_{APL1} \tag{7.12}$$

$$\mathbf{t}_{fl} = \mathbf{R}_{LAT}\mathbf{t}_{L1} + \mathbf{t}_{LAT} - \mathbf{R}_{LAT}\mathbf{R}^T_{AP}(\mathbf{R}_{AP}\mathbf{t}_{L1} + \mathbf{t}_{AP}) \tag{7.13}$$

$$\mathbf{t}_{fl} = \mathbf{t}_{LAT} - \mathbf{R}_{LAT}\mathbf{R}^T_{AP}\mathbf{t}_{AP} \tag{7.14}$$

However, equation (7.14) shows that the translation of the fluoroscopy set is dependent on the rotational matrices $\mathbf{R}_{LAT}$ and $\mathbf{R}_{AP}$. This can cause problems because $\mathbf{t}_{AP}$ is typically a large translation and so small fluctuations in the rotational parameters can cause large fluctuations in $\mathbf{t}_{fl}$. This can result in large variations in $\mathbf{t}_{fl}$ purely due to small rotational errors. To correct for this problem the rotational and translational matrices have been decoupled to as large a degree as possible by translating the final extrinsic parameters from both the AP and lateral views by the average AP translational parameter, $\overline{\mathbf{t}}_{AP}$. The average AP translational parameter is calculated by averaging the translations to all of the vertebrae in the AP view, for example, equation (7.15) shows how $\overline{\mathbf{t}}_{AP}$ is calculated for the case where there are two vertebrae.

$$\overline{t[k]}_{AP} = \frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}(t[k]_{APL1})_i + \frac{1}{N}\sum_{i=1}^{N}(t[k]_{APL2})_i\right) \qquad k = 1,2,3 \tag{7.15}$$

Where the index k represents the x, y and z components of the translation vector, $\mathbf{t} = (t[1], t[2], t[3])$ and there are N registrations to each vertebra. Using these translated parameters, the translation of the fluoroscopy set is now given by $\mathbf{t}'_{fl}$, see equation (7.16). Errors in the rotational matrices have a much smaller effect on $\mathbf{t}'_{fl}$ compared to $\mathbf{t}_{fl}$ because $\|\mathbf{t}_{AP} - \overline{\mathbf{t}}_{AP}\| \ll \|\mathbf{t}_{AP}\|$.

$$\mathbf{t}'_{fl} = \mathbf{t}_{LAT} - \overline{\mathbf{t}}_{AP} - \mathbf{R}_{LAT}\mathbf{R}^{T}_{AP}(\mathbf{t}_{AP} - \overline{\mathbf{t}}_{AP}) \tag{7.16}$$

The parameters defining the transformation of the fluoroscopy set, $F^k, k = 1,\ldots,6$ are calculated from $\mathbf{R}_{fl}$ and $\mathbf{t}'_{fl}$. The translational parameters, $F^k, k = 4,\ldots,6$, are obtained directly from the matrix $\mathbf{t}'_{fl}$. The rotational parameters, $F^k, k = 1,\ldots,3$ are calculated by decomposing $\mathbf{R}_{fl}$ into its constituent parameters using the method described in section 4.2.2.

7.3.1 Problems caused by combining in-plane and out-of-plane parameters

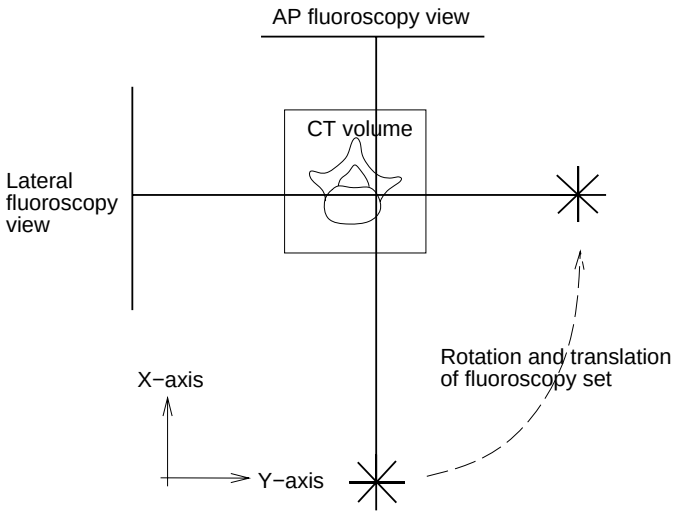
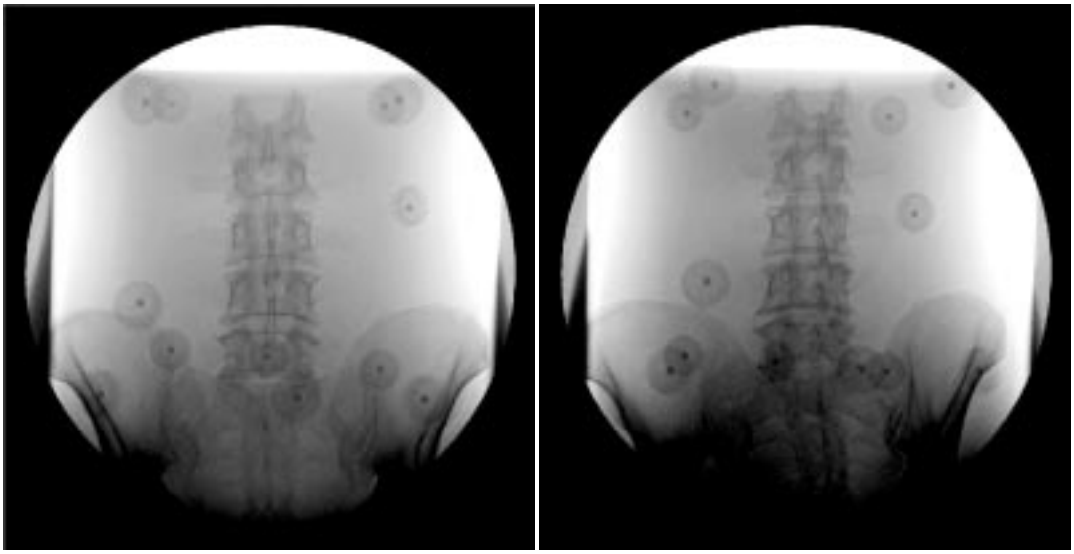


Figure 7.4: Showing how the in-plane translational parameter for the AP view (a Y translation) becomes the out-of-plane translational parameter in the lateral view and vice versa for an X translation. Note that translations along the z-axis (out of the page) remain in-plane for both the AP and lateral views.

As shown in the previous chapter there is a large difference in the expected accuracy of out-of-plane and in-plane translational parameters. A rotation of the fluoroscopy set alters the out-of-plane translation direction. Figure 7.4 shows the result of a 90° rotation of the fluoroscopy set. In the AP view the position of the CT volume along the Y-axis will be much more accurately determined than the position of the CT volume along the X-axis. The reverse will happen in the lateral view where the X direction will be much more accurately determined compared to the Y direction. Therefore, if the transformation of the fluoroscopy set is calculated from the AP to the lateral position, then there will be large errors in both the X and Y translational parameters. The



(a) (b)

Figure 7.5: Fluoroscopy images of spine phantom, anterior-posterior view (a) and view taken after rotating the fluoroscopy gantry by 5 degrees cranio-caudial rotation and 10 degrees oblique rotation (b).

size of the error in the Z direction (out of the page) will remain unaffected because this degree of freedom is in-plane in both the AP and lateral images. If the fluoroscopy set is rotated by less than 90° then the mixing of in-plane and out-of-plane errors will be reduced. If there is no rotation (only a translation) between the two fluoroscopy views then there should be no increase in the in-plane errors due to this effect.

7.4 A phantom experiment to test methodology

To investigate the validity of using $\mathbf{CV}'$ to estimate $\mathbf{RMSE}$ the following experiment was carried out using images of the lumbar spine phantom.

7.4.1 Method

The same fluoroscopy and CT images, as in chapter 6, were used. Also, a second fluoroscopy image was acquired. The transformation of the fluoroscopy set between the acquisition of the two fluoroscopy images was approximately a 5° cranio-caudial rotation (θ_y) and a 10° oblique rotation (θ_z). Both of the fluoroscopy images can be seen in figure 7.5. Because the transformation of the fluoroscopy set involved rotations in both θ_y and θ_z, a component of the out-of-plane translational error will be present in both of the in-plane translational errors calculated using *k-fold* cross validation i.e. the Y and Z components in $\mathbf{CV}'$.

The second fluoroscopy image was corrected for pincushion distortion (see section 4.4.2 for method) and a "gold-standard" registration was calculated using the fiducial markers. A ROI was defined for each of the vertebrae in the new fluoroscopy image. Sixty four registrations were carried out to each vertebra in each fluoroscopy image, using the same method as before to determine the starting positions (see section 6.2.1 for more details). The final extrinsic parameters were transformed from a coordinate system based on a reduced CT volume to a coordinate system based on the whole CT volume. The parameters were then filtered to remove failed registrations using the same method as before (see section 6.2.2). Then, for each vertebra, the successful final extrinsic parameters from the first view and the successful final extrinsic parameters from the second view were used to calculate the rotation and translation of the fluoroscopy set, $\mathbf{F}_{L1,\ldots,L5}$. Each possible permutation of first view extrinsic parameters and second view extrinsic parameters was carried out. This resulted in a maximum of 4096 sets of parameters defining the movement of the fluoroscopy gantry per vertebra. The *k-fold* cross validation technique was applied to these parameters and an estimated RMS error for $\mathbf{F}$ was calculated, $\mathbf{CV}$. This value represents the RMS error for two registrations (one to each fluoroscopy image), and so $\mathbf{CV}$ was multiplied by $1/\sqrt{2}$ to estimate the RMS error for a single registration, $\mathbf{CV}'$.

The real RMS errors, for registrations to each fluoroscopy image, were also calculated using the "gold-standard" registration. These were then combined as shown by equation (7.17) to give a total RMS error, $\mathbf{RMSE}_{tot}$. This experiment was repeated seven times, using a different similarity measure each time.

$$RMSE^k_{tot} = \sqrt{\frac{1}{2}\left((RMSE^k_{image1})^2 + (RMSE^k_{image2})^2\right)}, \quad k = 1, \ldots, 6 \tag{7.17}$$

7.4.2 Results

Table 7.1 shows the values of the two error measures $\mathbf{CV}'$ and $\mathbf{RMSE}_{tot}$, and the failure rate for each similarity measure. The mean difference, $\frac{1}{7}\Sigma(\mathbf{CV}' - \mathbf{RMSE}_{tot})$, and the mean fractional difference, $\frac{1}{7}\Sigma((\mathbf{CV}' - \mathbf{RMSE}_{tot})/\mathbf{RMSE}_{tot})$, between the two error measures are shown in table 7.2.

7.4.3 Discussion

The results show that $\mathbf{CV}'$ has a tendency to underestimate the RMS error as measured by $\mathbf{RMSE}_{tot}$. The average size of this underestimation is shown in Table 7.2. For five of the six

Similarity measure	Type	Error measure	Extrinsic parameters						No. fail (%)
			rotation (degrees)			translation (mm)			
			θ_x	θ_y	θ_z	X	Y	Z	
Cross correlation	2	$\mathbf{CV}'$	0.45	0.97	0.57	10.2	1.32	0.63	18.4
		$\mathbf{RMSE}_{tot}$	0.57	1.66	0.72	10.6	1.37	1.26	
Entropy	2	$\mathbf{CV}'$	0.27	0.51	0.51	9.0	0.93	0.46	15.6
		$\mathbf{RMSE}_{tot}$	0.36	0.94	0.56	12.4	1.14	1.21	
Mutual information	2	$\mathbf{CV}'$	0.86	1.51	1.00	20.4	2.06	1.10	41.1
		$\mathbf{RMSE}_{tot}$	0.98	1.74	1.39	19.4	1.82	1.66	
Gradient correlation	3	$\mathbf{CV}'$	0.25	0.45	0.31	3.1	0.56	0.23	2.8
		$\mathbf{RMSE}_{tot}$	0.34	0.40	0.27	5.3	0.57	0.48	
Pattern intensity	1	$\mathbf{CV}'$	0.25	0.37	0.34	3.8	0.49	0.22	0.3
		$\mathbf{RMSE}_{tot}$	0.37	0.41	0.29	6.1	0.63	0.46	
Gradient difference	3	$\mathbf{CV}'$	0.26	0.34	0.37	4.2	0.44	0.25	0.2
		$\mathbf{RMSE}_{tot}$	0.36	0.34	0.31	6.9	0.63	0.50	
Local correlation	1	$\mathbf{CV}'$	0.28	0.35	0.36	4.1	0.76	0.25	1.4
		$\mathbf{RMSE}_{tot}$	0.34	0.40	0.34	4.9	0.74	0.49	

Table 7.1: Comparing two methods of calculating registration accuracy. Accuracy calculated using an independent "gold-standard" $\mathbf{RMSE}_{tot}$, and accuracy calculated using a *k-fold* cross validation technique and two fluoroscopy images, $\mathbf{CV}'$.

	Extrinsic parameters					
	θ_x	θ_y	θ_z	X	Y	Z
Mean difference	-0.10	-0.20	-0.07	-1.37	-0.03	-0.41
Mean fractional difference	-0.23	-0.16	-0.02	-0.23	-0.09	-0.50

Table 7.2: The mean difference and the mean fractional difference between the values of the two error measures, $\mathbf{CV}'$ and $\mathbf{RMSE}_{tot}$, shown in table 7.1.

parameters the mean fractional difference is between -0.02 and -0.23. Therefore, for these parameters the mean underestimation of the registration error is less than one quarter of the $\mathbf{RMSE}_{tot}$ error value. The exception is the error value of the Z translation parameter, where the underestimation was half the value of $\mathbf{RMSE}_{tot}$. To investigate the cause of this much larger underestimation the registration results to each fluoroscopy image were analysed separately. The result of this investigation showed that there was a constant bias of approximately 0.4mm between the final registration positions and the "gold-standard" registration in the Y and Z parameters for both fluoroscopy images. The bias in the Y direction was not detected in the results shown in tables 7.1 and 7.2 because the Y translation direction contained an out-of-plane component due to the 10 degree oblique rotation between the two fluoroscopy views. The origin of this bias was investigated by producing DRRs at the "gold-standard" position and at the "gold-standard" position plus the in-plane bias. Visual inspection was used to investigate which DRR was registered more accurately

to the fluoroscopy images. Although it is difficult to observe such small misregistrations visually, the DRR produced at the "gold-standard" position appeared to be more accurately registered with regard to the position of the fiducial markers. However, the DRR produced at the "gold-standard" position plus the in-plane bias appeared to be more accurately registered to the vertebrae. Therefore, this in-plane bias may be due to a residual uncorrected image distortion violating the rigid body assumption, rather than a misregistration.

I conclude from this phantom experiment that the estimated RMS error using *k-fold* cross validation $\mathbf{CV}'$, is a reasonable approximation to the average RMS error measured using an independent "gold-standard", $\mathbf{RMSE}_{tot}$. The estimated RMS error has a tendency to underestimate the registration errors, however, the size of this underestimation is relatively small compared to the size of the errors in the rotational and out-of-plane translational parameters, less than a quarter of $\mathbf{RMSE}_{tot}$. This underestimation should be kept in mind though when analysing the results in the following sections. A problem was found when analysing the in-plane translational parameters due to a constant bias between the final registration positions and the "gold-standard" registration position. However, this bias is believed to be due to problems in the calculation of the "gold-standard" registration rather than problems in the estimated RMS error measure.

7.5 Comparison using clinical data

Four clinical datasets from aortic stenting procedures were used, each of which consisted of two fluoroscopy images and a CT scan of the lumbar vertebrae. The fluoroscopy images were corrected for distortion using the method described in section 4.4.2. The movement of the fluoroscopy set, between when the two images were acquired, varied from just an in-plane translation to a 20-25° oblique rotation. In two of the datasets, four of the lumbar vertebrae were visible in both of the fluoroscopy images and in the CT volume. In the other two datasets only three lumbar vertebrae were visible in all three images.

7.5.1 Method

Regions of interest were manually defined around each of the vertebrae in the fluoroscopy images, and reduced CT volumes were created for each vertebra in the CT images.

A starting position for the registration algorithm was found in the following way. The distance between the centre of the CT volume and the x-ray source was estimated to be 75% of the focal distance. The rotation about θ_z (axis normal to the slice plane) was estimated from the

fluoroscopy image where a -90° rotation was used for a PA view. For an oblique view visual inspection was used, comparing DRRs produced using different values of θ_z with the fluoroscopy image. The other two rotational parameters θ_x and θ_y were set to zero. The in-plane translational parameters were estimated by manually picking corresponding points on the fluoroscopy image and the DRR and then overlaying these points by altering the Y and Z parameters. The algorithm was then run from this starting position using the gradient difference similarity measure. The final extrinsic parameters from this registration were used as a mean starting position. Sixty four individual starting positions were calculated by perturbing the mean position by $\pm\Delta\mathbf{E}$, see table 6.2. There is a possibility that, because the gradient difference measure is used to calculate the mean position, the experiment is biased in favour of the gradient difference measure. However, $\Delta\mathbf{E}$ is much larger than the expected accuracy of the algorithm and so any bias caused by using the gradient difference measure to obtain the mean position should have a negligible effect on the final registration results.

The single point picking exercise to overlay corresponding features from the DRR and fluoroscopy image (as described in section 6.2.1) was also repeated. But in the case of clinical images the mean starting position was used instead of the "gold-standard" position (which does not exist). Also, because the clinical fluoroscopy images used were of size 512×512 pixels compared to the phantom fluoroscopy image which was of size 1024×1024 pixels, the Gaussian distributed error (used to simulate observer error) was of standard deviation 3 pixels (compared to 6 pixels for the phantom images).

Sixty four registrations were carried out to each vertebra in each fluoroscopy image. Failed registrations were removed, as described in section 6.2.2, however, the mean starting position was used instead of the "gold-standard" position. The final extrinsic parameters were transformed to represent a rotation about the entire CT volume rather than around a reduced CT volume.

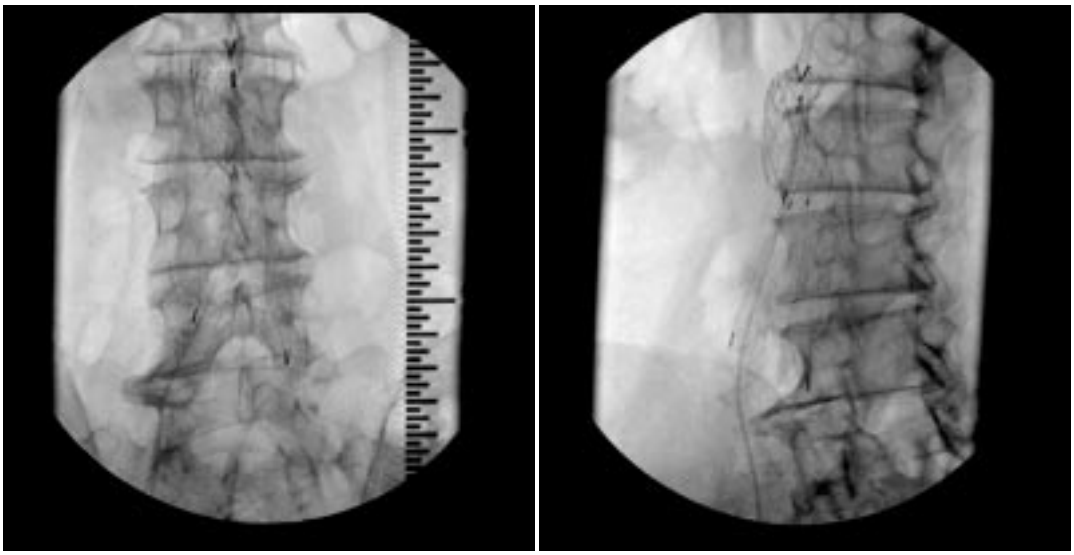
The registration results to each vertebra were used in turn to calculate a set of transformations of the fluoroscopy gantry $\mathbf{F}_i$. These sets of transformations were then used to calculate $\mathbf{CV}'$, as described in section 7.2.

The datasets and the registration results are presented in the following four sections. The ordering of the datasets was based on my initial perception of which would be the easiest and hardest to register to, beginning with the easiest dataset. In each case, initially the images are described, and then the estimated RMS error, $\mathbf{CV}'$, and the failure rate for each similarity measure are presented. The mean final registration positions to each vertebra $\overline{\mathbf{E}}_{PA}$, when registering to a PA fluoroscopy image are also shown. The values of $\overline{\mathbf{E}}_{PA}$ were calculated using the most accurate

similarity measure for each dataset. By comparing the values of $\overline{\mathbf{E}}_{PA}$ between different vertebrae the amount of spinal deformation between the pre-operative CT image and the intra-operative fluoroscopy image can be calculated.

7.5.2 Dataset 1

The fluoroscopy images in dataset 1 are shown in figure 7.6. The fluoroscopy images were of size 512×512 pixels and were taken with a 25cm field of view. The transformation between the fluoroscopy views was approximately a 25° oblique rotation. The fluoroscopy images show a deployed aortic stent and a guidewire. Fluctuations due to soft tissue structures are also visible. The CT scan voxel size was $0.938\times0.938\times3.0$mm and the image dimensions were $320\times320\times61$ voxels. Registrations were carried out to the L3, L4 and L5 vertebrae.



(a) (b)

Figure 7.6: Distortion corrected PA (a) and oblique (b) fluoroscopy images from dataset 1. Taken on a Philips MD 3 with a 25cm field of view. Registrations carried out on vertebrae L3, L4 and L5.

Results from these registrations are shown in table 7.3. The results from all of the similarity measures were reasonably accurate. The type '2' measures, see table 5.1, can be seen to be the least accurate and mutual information was the worst. The most accurate measure is pattern intensity with a RMS rotational error of just 0.43°. Due to the 25° rotation of the fluoroscopy set a component of the Y translation direction is out-of-plane. This is the cause of the large errors in the Y translation direction for most of the measures. Two of the similarity measures, pattern intensity and gradient difference, were very accurate in both the in-plane and out-of-plane translational

parameters, both have errors below 0.9mm in the Y translation direction, despite this direction being partially out-of-plane.

The mean final extrinsic parameters for registrations to each vertebrae in the PA image, using the pattern intensity similarity measure, are shown in table 7.4. The difference between the mean final extrinsic parameters for different vertebra is due to spinal deformation. The amount of spinal deformation can be seen to be small, an average of 0.85° and 1.46mm, though the difference between the mean values are significant (95% confidence level). The size of the deformation is large compared to the estimated RMS error measurements shown in table 7.3. This shows the importance of using a validation scheme which accounts for spinal deformation. If the method used had not been invariant to spinal deformation then the measured RMS values would have been greatly increased purely due to spinal deformation.

Similarity measure	Type	$\mathbf{CV}'$						No. fail, (%)
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	
Cross correlation	2 B	0.80	0.84	0.42	4.7	2.04	0.84	2.9
Entropy	2 A	0.91	0.84	0.28	5.0	2.79	1.06	1.0
Mutual Information	2 A	0.80	1.32	0.91	8.6	3.43	1.32	17.2
Gradient correlation	3 B	0.55	0.69	0.69	4.6	1.82	0.41	12.0
Pattern Intensity	1 A	0.41	0.58	0.21	2.5	0.89	0.58	6.5
Gradient difference	3 A	0.76	0.56	0.73	2.8	0.76	0.70	4.9
Local correlation	1 A	0.44	0.52	0.39	4.8	1.37	0.57	8.6

Table 7.3: Estimated RMS error values from dataset 1. Based on the movement of the fluoroscopy gantry calculated from registrations to the L3, L4 and L5 vertebrae.

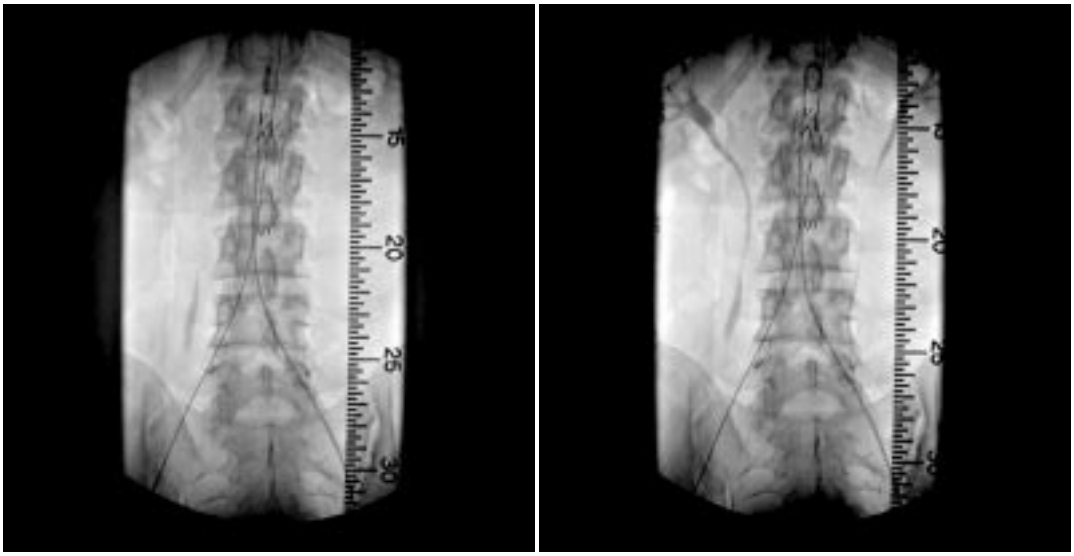
Vertebra	$\overline{\mathbf{E}}_{PA}$					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
L3	0.73	4.53	-95.31	844.8	15.80	68.28
L4	-0.41	5.42	-95.02	841.7	15.52	70.09
L5	0.45	3.63	-94.41	843.4	14.39	70.35

Table 7.4: Average final registration positions to each vertebra using the PA fluoroscopy image and pattern intensity similarity measure.

7.5.3 Dataset 2

The fluoroscopy images in dataset 2 are shown in figure 7.7. The fluoroscopy images were of size 512×512 pixels and were taken with a field of view of 38cm. The transformation between fluoroscopy views was approximately an in-plane translation. The images show a deployed aortic

stent and two guidewires, one in each femoral artery. There is a small amount of intensity variation caused by soft tissue structures and some contrast medium is visible, being filtered by the kidneys in image 7.7(b). The CT scan voxel size was 0.586×0.586×3.0mm and the image dimensions were 512×512×61 voxels. Registrations were carried out to the L2, L3, L4 and L5 vertebrae.



(a) (b)

Figure 7.7: Distortion corrected PA fluoroscopy images from dataset 2. Note that there has been an in-plane translation between when the two views were acquired. Taken on a Philips MD 3 with a 38cm field of view. Registrations carried out on vertebrae L2, L3, L4 and L5.

Results from these registrations are shown in table 7.5. Again the type '2' measures were the least accurate. The other four measures were again accurate and some were extremely robust, pattern intensity yielding only 2 failures from 512 registrations. The largest rotational error from the type '1' and '3' measures was 0.79° and the largest in-plane translational error was 0.95mm.

Similarity measure	Type	$\mathbf{CV}'$						No. fail, (%)
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	
Cross correlation	2 B	1.08	0.42	0.48	5.8	1.10	0.57	6.1
Entropy	2 A	0.89	0.47	0.63	8.6	1.32	0.94	4.1
Mutual Information	2 A	1.25	1.45	1.22	9.6	1.59	0.94	46.5
Gradient correlation	3 B	0.58	0.22	0.46	2.6	0.72	0.25	4.5
Pattern Intensity	1 A	0.79	0.37	0.52	2.9	0.95	0.30	0.4
Gradient difference	3 A	0.61	0.25	0.50	3.0	0.84	0.26	0.8
Local correlation	1 A	0.61	0.42	0.47	3.7	0.84	0.36	2.3

Table 7.5: Estimated RMS error values from dataset 2. Based on the movement of the fluoroscopy gantry calculated from registrations to the L2, L3, L4 and L5 vertebrae.

Table 7.6 shows the mean final registration positions to each vertebra in figure 7.7(a) using the gradient correlation similarity measure. The final extrinsic parameters in table 7.6 show that the range of deformation in all the extrinsic parameters is larger than 1mm or degree. The largest amount of rotational deformation occurred due to flexion-extension of the spine (θ_y direction) with a range of 3.4° over the four vertebrae.

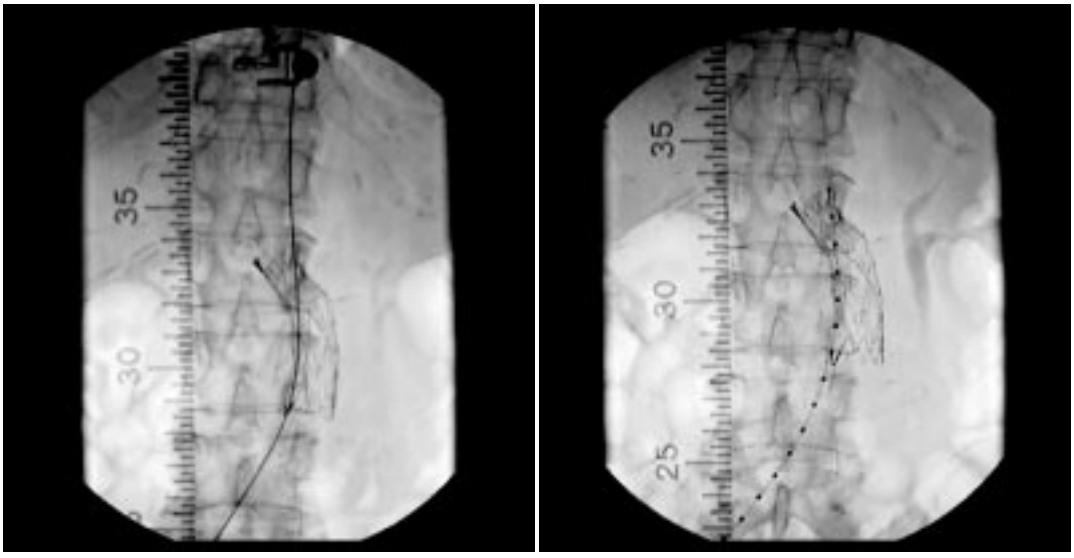
Vertebra	$\overline{\mathbf{E}}_{PA}$					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
L2	1.47	-2.18	-80.95	812.9	-54.67	0.29
L3	2.22	-1.90	-81.04	808.5	-53.91	-0.36
L4	1.02	-0.87	-80.65	810.1	-53.95	-0.82
L5	1.54	1.29	-79.36	811.6	-55.25	-1.46

Table 7.6: Average final registration positions to each vertebra in figure 7.7(a) using the gradient correlation similarity measure.

7.5.4 Dataset 3

The fluoroscopy images in dataset 3 are shown in figure 7.8. The fluoroscopy images were of size 512×512 pixels and were taken with a 25cm field of view. The transformation between fluoroscopy views was approximately an in-plane translation. The fluoroscopy images show a deployed aortic stent, a guidewire in (a) and a measurement catheter in (b) and also the measurement ruler now underlies the right hand side of the vertebral body in the fluoroscopy image. There has also been an increase in the intensity fluctuations caused by soft tissue structures compared to datasets 1 and 2. The CT scan voxel size was 1.094×1.094×3.0mm and the image dimensions were 320×320×59 voxels. Registrations were carried out to the L1, L2, L3 and L4 vertebrae.

Results from these registrations are shown in table 7.7. This dataset proved to be the worst in terms of robustness of the similarity measures. The most robust measure, local correlation, had a 21.3% failure rate. Failure rates were all high for the type 2 measures, especially cross correlation (82.0%) and mutual information (98.2%). The cause of the high failure rate was investigated by analysing the position of the failed registrations when using the pattern intensity and gradient difference similarity measures. A large number of the failed final registration positions appeared to be random, though some of the failed registrations clustered into a local minimum which is believed to be caused by the presence of the ruler underlying the structure from the vertebrae. The position of this local minimum involved a translation of the CT volume towards the x-ray source and to the left of the image. This increased the amount of overlap between high pixel values in the



(a) (b)

Figure 7.8: Distortion corrected PA fluoroscopy images from dataset 3. Note that there has been an in-plane translation between when the two views were acquired. Taken on a Philips MD 3 with a 25cm field of view. Registrations carried out on vertebrae L1, L2, L3 and L4.

DRR and the pixels containing the ruler.

Despite the high failure rate all of the type '1' and '3' measures were fairly accurate, with estimated RMS errors all below 1° and in-plane translational errors below 1.5mm.

Similarity measure	Type	$\mathbf{CV}'$						No. fail, (%)
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	
Cross correlation	2 B	0.50	0.71	0.31	7.5	0.77	0.39	82.0
Entropy	2 A	1.77	0.79	0.56	14.1	0.26	0.95	58.6
Mutual Information	2 A	1.92	2.46	4.71	48.7	5.60	2.15	98.2
Gradient correlation	3 B	0.43	0.96	0.79	10.1	1.17	0.94	32.8
Pattern Intensity	1 A	0.54	0.58	0.90	14.0	1.49	0.97	27.5
Gradient difference	3 A	0.46	0.72	0.96	8.0	0.83	0.75	28.5
Local correlation	1 A	0.59	0.61	0.83	9.8	1.15	0.67	21.3

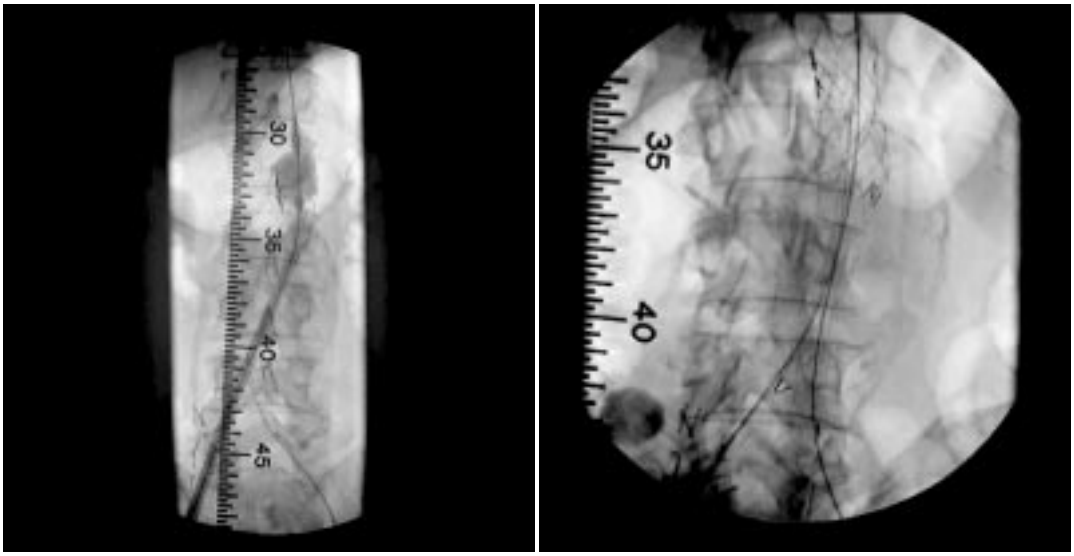
Table 7.7: Estimated RMS error values from dataset 3. Based on the movement of the fluoroscopy gantry calculated from registrations to the L1, L2, L3 and L4 vertebrae.

Table 7.8 shows the mean final registration positions to each vertebra in figure 7.8(a) using the gradient difference similarity measure. This dataset showed the largest amounts of deformation with a range of over 5° in flexion-extension deformation and a stretching of 5mm in the Z direction.

Vertebra	$\overline{\mathbf{E}}_{PA}$					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
L1	0.63	-2.14	-94.19	841.6	40.24	12.38
L2	1.04	-1.12	-94.39	846.6	41.09	11.42
L3	0.75	-0.14	-94.00	850.4	40.97	9.05
L4	0.90	3.18	-93.90	846.5	40.56	7.15

Table 7.8: Average final registration positions to each vertebra in figure 7.8(a) using the gradient difference similarity measure.

7.5.5 Dataset 4



(a) (b)

Figure 7.9: Distortion corrected PA (a) and oblique (b) fluoroscopy images from dataset 4. Taken on a Philips MD 3 with a 38cm field of view for PA view and 25cm field of view for oblique view. Registrations carried out on vertebrae L2, L3 and L4.

The fluoroscopy images in dataset 4 are shown in figure 7.9. The fluoroscopy images were of size 512×512 pixels. The PA view has been taken with a 38cm field of view and the oblique view has been taken at a higher magnification, with a 25cm field of view. The transformation between the fluoroscopy views was approximately a 20° oblique rotation. These fluoroscopy images show the most complicated scenes and so were expected to be the most difficult to register to. The PA view shows a partially deployed stent with contrast in the expanding balloon. The ruler also underlies structure from the vertebrae. The oblique view shows a deployed stent and two guidewires. As in dataset 3 there are some large intensity fluctuations caused by soft tissue structures. The CT scan voxel size was 0.586×0.586×3.0mm and the image dimensions were

512×512×52 voxels. Registrations were carried out to the L2, L3 and L4 vertebrae.

The registration results are shown in table 7.9. Note that no successful registrations were achieved by the cross correlation similarity measure registering to the PA view of L2 or to the oblique view of L3, therefore, it was not possible to calculate $\mathbf{CV}'$ for the cross correlation measure. High failure rates and low accuracy were observed for the other type '2' measures. All of the similarity measures have a large Y translational error due to the rotation of the fluoroscopy set causing this parameter to be partially out-of-plane. The in-plane errors (Z direction) for pattern intensity and gradient difference are small, both below 0.66mm. The most accurate and robust measure was pattern intensity.

Similarity measure	Type	$\mathbf{CV}'$						No. fail, (%)
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	
Cross correlation	2 B	-	-	-	-	-	-	47.1
Entropy	2 A	0.96	3.18	1.69	7.4	6.04	0.88	30.5
Mutual Information	2 A	4.03	5.55	3.36	14.1	5.41	6.92	69.0
Gradient correlation	3 B	1.40	0.51	1.35	21.7	7.48	1.29	10.9
Pattern Intensity	1 A	0.64	0.62	0.83	12.1	3.57	0.60	5.2
Gradient difference	3 A	1.09	0.68	1.08	15.0	4.29	0.66	8.3
Local correlation	1 A	0.97	0.84	1.60	24.3	4.40	1.43	7.6

Table 7.9: Estimated RMS error values from dataset 4. Based on the movement of the fluoroscopy gantry calculated from registrations to the L2, L3 and L4 vertebrae. Note that no successful registrations occurred using the cross correlation measure registering to the PA view of L2 and the oblique view of L3, therefore, no measurements of registration error were possible using *k-fold* cross validation.

The mean final registration positions to each vertebra using the PA fluoroscopy image and the pattern intensity similarity measure are shown in table 7.10. The final extrinsic parameters show a small amount of deformation, in the order of 2mm or degrees for most of the parameters.

Vertebra	$\overline{\mathbf{E}}_{PA}$					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
L2	-0.31	-4.18	-85.43	854.9	-52.70	-24.30
L3	1.13	-5.79	-83.99	852.2	-53.52	-27.23
L4	1.04	-3.33	-82.74	861.5	-54.73	-27.97

Table 7.10: Average final registration positions to each vertebra using the PA fluoroscopy image and pattern intensity similarity measure.

7.5.6 Discussion

The results in sections 7.5.2 to 7.5.5 show that the type '2' measures were consistently less accurate and less robust (except to dataset 1) than the type '1' and '3' measures. These results back up my earlier findings using phantom data, where the accuracy and robustness of the type '2' measures were affected by the presence of soft tissue structures. The other four similarity measures all produced accurate registrations, an overall estimated RMS rotational error of 0.76° and an average RMS in-plane translational error of 0.85mm[1]. The type '1' and '3' measures were also fairly robust, with an overall average failure rate of 10% for pattern intensity, gradient difference and local correlation and 15% for gradient correlation. Overall, the pattern intensity and gradient difference similarity measures appear slightly more accurate, this also backs up my findings from chapter 6.

The majority of the failures for the type 1 and 3 measures occurred when dataset '3' was used. The high failure rate for data set 3 shows that under certain circumstances differences between the images (i.e. the presence of the interventional ruler) can create local minimum. Despite the high failure rate the the type 1 and 3 measures were still accurate when registering to data set 3. For use in a clinical setting any failure rate greater than zero could be dangerous if the misregistrations are not detected. Therefore, it is imperative that, if the algorithm is used in a clinical setting, a reliable system must be in place to detect misregistrations. Such systems are discussed in section 9.5.2 which outlines future work. The major implications of such a high failure rate are that it would reduce confidence in the algorithm, increase the chance of a misregistration remaining undetected and increase the registration time.

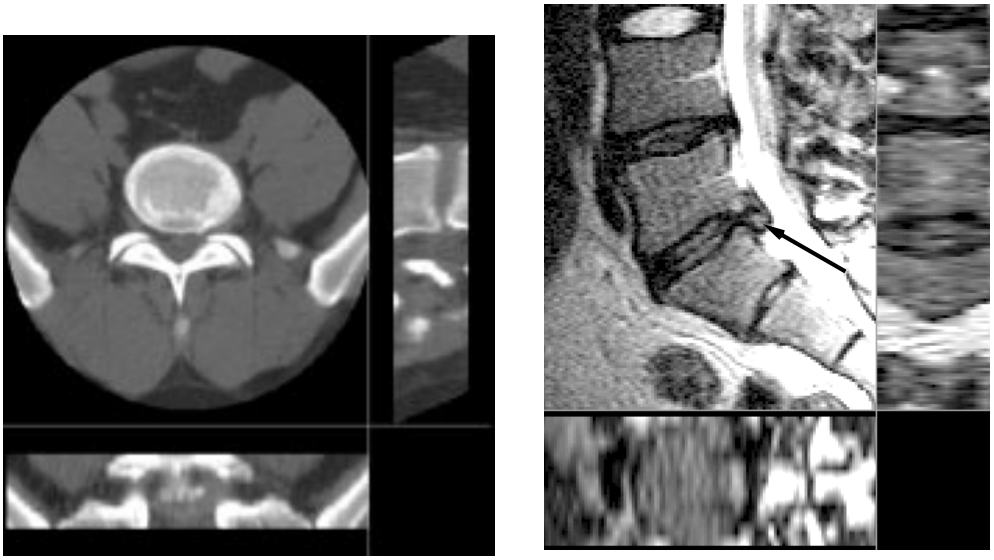
Using the mean final registration positions per vertebra the amount of spinal deformation can be calculated. The amount of deformation observed ranged from zero up to 5.3°. The largest amount of deformation was observed in the θ_y rotation direction, which corresponds to a flexion-extension of the spine and in the Z translation direction which represents a slight lengthening/shortening of the spine. The values of rotational spinal deformation are easily within the range of movement of the lumbar spine [36, 56]. The size of the deformation was large when compared with the estimated RMS errors. This shows that it was necessary to use the two view approach to remove the effect of spinal deformation. It also validates the assumption that the

[1]The average RMS in-plane translational error was calculated using the errors in the Y and Z direction from datasets 2 and 3 and the errors in just the Z direction for datasets 1 and 4. The errors in the Y direction were not included for datasets 1 and 4 because the rotation of the fluoroscopy set makes a component of the Y direction out-of-plane for these datasets.

deformation occurs between the pre-operative and intra-operative images and there is very little spinal deformation between the two fluoroscopy images.

7.6 2D-3D registration using images from a percutaneous laser nucleotomy procedure

This section describes a registration carried out between a fluoroscopy image taken on the BV29 mobile fluoroscopy set and a pre-operative CT scan of a patient who underwent a percutaneous laser nucleotomy (PLN) procedure. Pre-operative CT and MR scans were acquired for diagnosis and for planning the procedure, these images are shown in figure 7.10.



(a) (b)

Figure 7.10: Reformatted axial, sagittal and coronal slices from the pre-operative CT volume (a) and the pre-operative MR volume (b) of a patient who underwent a percutaneous laser nucleotomy procedure. The black arrow shows the position of the prolapse in the sagittal MR slice.

7.6.1 Method

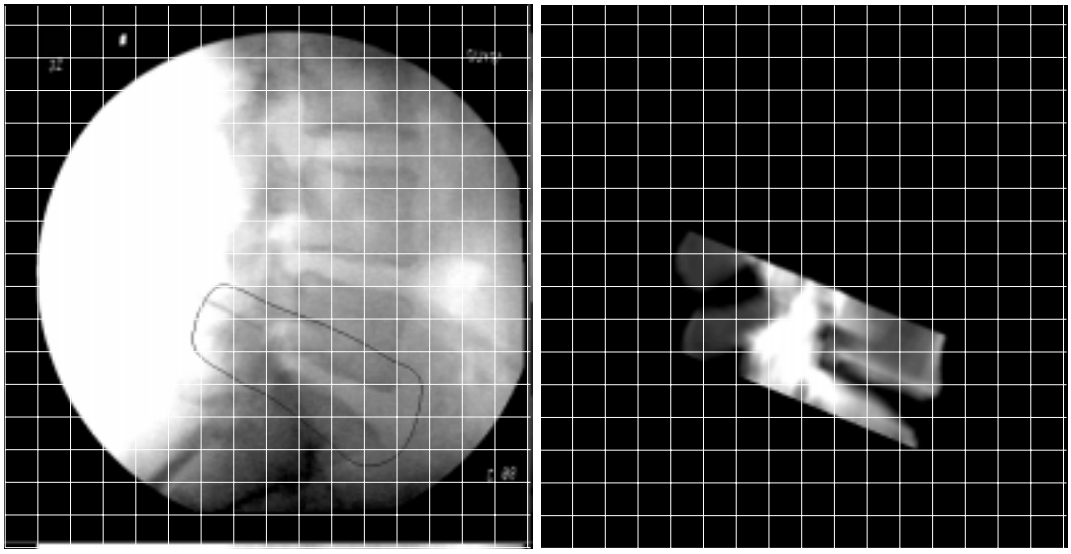
The CT acquisition protocol for this procedure was to obtain a small CT volume for each vertebral disc, where the CT gantry was tilted so that the slice plane was parallel to each disc. This resulted in a number of small CT volumes, each having a different gantry tilt angle. Consequently there were only a small number of CT slices to work with, 13 slices spanning the inferior half of the L5 vertebra, the L5/S1 disc and the superior surface of the S1 vertebra. Because of the limited amount of data the registration was carried out to both the L5 and S1 vertebrae, even though this

meant registering on two rigid bodies which may have moved with respect to each other.

The pre-operative CT volume was registered to the intra-operative fluoroscopy image shown in figure 7.11(a). A reduced CT volume, which comprised of a box surrounding the L5 and S1 vertebrae was segmented from the pre-operative CT image. A ROI was manually drawn in the fluoroscopy image to contain the inferior half of the L5 vertebra and superior half of the S1 vertebra, this ROI is shown as the black outline in figure 7.11(a). A starting estimate was found using the same method as described earlier, see section 7.5. Though the in-plane rotation angle θ_x, also needed to be adjusted until a suitable position was found. This was achieved using visual inspection.

The pattern intensity and gradient difference measures were used to register the reduced CT volume to the fluoroscopy image. These two similarity measures were chosen as on average they were the most accurate and robust in previous experiments. Sixty four registrations were carried out using each measure and the starting positions were calculated using the same method as before.

The pre-operative CT volume was also registered to the pre-operative MR volume using an algorithm designed by Studholme *et al.* using normalised mutual information [112, 116]. This registration, in conjunction with the mean final extrinsic parameters from the 2D-3D registrations, were used to project information from the pre-operative MR volume into the intra-operative fluoroscopy image.



(a) (b)

Figure 7.11: Intra-operative fluoroscopy from Philips BV29 (a) the black outline shows the ROI in the fluoroscopy image used for registration. A DRR (b) produced at the average final registration position using the pattern intensity similarity measure. The overlaid grid is included so that the spatial position of features in both images can be compared.

7.6.2 Results

The registration results are presented in tables 7.11 and 7.12. Table 7.11 shows the number of failed registrations and the standard deviation of each of the extrinsic parameters for the successful registrations. Table 7.12 shows the mean final registration positions for both similarity measures. The results for gradient difference in table 7.11 are shown twice. The results labelled gradient difference 1 (and pattern intensity) are calculated using the standard criteria to determine whether a registration was a failure, i.e. $|E_i^k - E_{GS}^k| > \Delta E^k, k = 1 \rightarrow 6$ (equation 6.2), where $\mathbf{E}_{GS}$ in this case was the mean starting position. The results when using the gradient difference measure showed a lower failure rate than pattern intensity. However, the final extrinsic parameters when using the gradient difference measure had a much larger standard deviation than the results using pattern intensity. The cause of this larger standard deviation was investigated and both measures were found have a local minimum in a very similar position. However, the position of the local minima found using the gradient difference measure was just inside the boundary which defines a successful registration, while the local minima found using the pattern intensity measure was just outside the boundary. Because of this a second set of results was calculated, gradient difference 2, which used a slightly more stringent criteria for successful registrations, $|E_i^k - E_{GS}^k| > 0.9 \times \Delta E^k, k = 1 \rightarrow 6$.

Similarity measure	Type	Standard deviation						No. fail (%)
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	
Pattern Intensity	1	0.25	0.16	0.16	3.9	0.15	0.14	25.0
Gradient difference 1	3	0.55	1.29	2.80	9.2	0.31	0.22	10.9
Gradient difference 2	3	0.43	0.15	0.30	9.1	0.24	0.22	26.6

Table 7.11: Standard deviations of registration results using images from a percutaneous laser nucleotomy procedure. The results labelled pattern intensity and gradient difference 1 were calculated using the standard method to remove failed registrations (see equation 6.2) where $\mathbf{E}_{GS}$ was replaced by the average starting position. The results labelled gradient difference 2 use a slightly stricter criteria for successful registrations where $|E_i^k - E_{GS}^k| > 0.9 \times \Delta E^k, k = 1 \rightarrow 6$, see the text for the explanation behind using this stricter criteria.

Similarity measure	Type	mean final extrinsic parameters					
		rotation (degrees)			translation (mm)		
		θ_x	θ_y	θ_z	X	Y	Z
Pattern Intensity	1	-22.34	0.72	-182.51	691.6	-12.50	4.51
Gradient difference 2	3	-22.11	0.24	-182.60	692.1	-12.52	4.44

Table 7.12: The mean values of the successful final registration parameters.

The gradient difference 2 results and the results for pattern intensity show the two measures to have a similar robustness. Little can be said about the accuracy of the registrations as there is no independent "gold-standard" and the *k-fold* cross validation approach cannot be used due to the small size of the CT volume. Though the final mean parameters from both measures can be seen from table 7.12 to be very similar and DRRs produced using these parameters appear to be accurately registered using visual inspection. Figure 7.11(b) shows a DRR produced using the reduced CT volume at the mean registration position using pattern intensity.

Figure 7.12 shows the result of projecting information from a pre-operative MR image onto the intra-operative fluoroscopy image. The nucleus pulpose and prolapse from the L5/S1 disc were segmented from the MR image. This segmented image was firstly transformed into the coordinate system of the CT volume (using the 3D-3D registration result) and then projected onto the fluoroscopy image using the mean final 2D-3D registration parameters.

Figure 7.12: Intra-operative fluoroscopy image showing a projection of the nucleus pulpose plus prolapsed material segmented from the pre-operative MR image.

7.6.3 Discussion

The results from registrations using the PLD images show that the algorithm can register with high precision using only a small number of CT slices. As there is no "gold-standard" it is difficult to assess the registration accuracy. DRRs produced at the mean final positions, see figure 7.11(b), appear to be accurately aligned using visual inspection. Also, the fact that both similarity measures obtained very similar final registration positions indicates that the algorithm has found the correct

solution rather than a local minimum. The failure rate for the similarity measures were fairly high, 25% and 26.6%. The cause of the high failure rate was found to be a local minimum, the position of which is shown in table 7.13. This local minimum was responsible for nearly all the unsuccessful registrations: all apart from one registration using the gradient difference measure. Comparing the position of the local minimum in table 7.13 with the mean final registration position in table 7.12 shows that the main difference is a rotation about θ_z. The suspected cause of this local minimum is illustrated in figure 7.13. It is believed to be due to the sagittal line of symmetry through a vertebra. If a vertebra was perfectly symmetrical about a sagittal line and if parallel projection is used, then DRRs produced at the two positions shown in figure 7.13 would be identical. Though due to perspective projection and imperfections in the symmetry of vertebra the two DRRs will actually be slightly different. Therefore, to aid the registration algorithm a short focal distance would be preferred, however, to decrease blurring and shape distortion (due to different magnification effects) in the fluorosocopy image a long focal distance is preferred [41].

Similarity measure	Type	Mean extrinsic parameters from local minima						No. regns
		rotation (degrees)			translation (mm)			
		θ_x	θ_y	θ_z	X	Y	Z	(%)
Pattern Intensity	1	-22.83	-2.58	-174.57	688.9	-12.20	4.59	25.0
Gradient difference 2	3	-22.84	-3.18	-175.24	691.7	-12.05	4.52	25.0

Table 7.13: Mean of the unsuccessful final registration positions which found a large local minimum. Note the No. regns column shows the percentage of the total registrations which found the local minima.

Lines of symmetry causing local minimum can be a problem for 2D-3D registration algorithms, especially as DRRs produced at the position of the local minimum, can appear to be accurately registered using visual inspection. The local minimum found in this experiment probably has a particularly large effect due to the small amount of data used, only 13 slices. Solutions to coping with such local minimum could be to use a different search strategy which is able to climb out of local minimum, or to accept that such lines of symmetry and subsequent local minimum can cause problems and do not attempt to register to fluoroscopy images where they may create problems, i.e. use an oblique rather than a lateral view.

Figure 7.12 shows a contoured outline of the projection of the nucleus pulpose and prolapse on the intra-operative fluoroscopy image. Initially the outline of the nucleus pulpose was thought to be too large because it overlapped the edges of the vertebrae. Though upon further investigation the overlap between the outline of the projection of the nucleus pulpose and the vertebral bodies is believed to be due to the concave nature of the superior and inferior surfaces of the lumbar vertebrae.

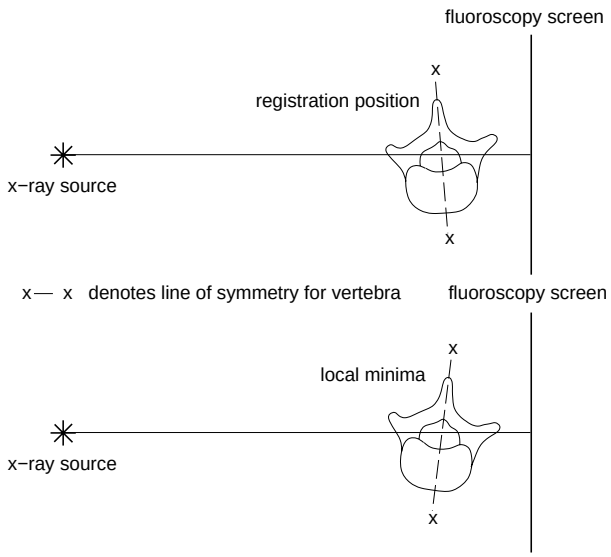


Figure 7.13: Showing the position of the local minima relative to the registration position. The local minima is caused due to the sagittal line of symmetry, line XX, through a vertebra. In the registration position the line of symmetry is inclined so the vertebral body points towards the fluoroscopy screen. The local minima is formed when the line of symmetry is inclined at the same angle, but this time the vertebral body points away from the fluoroscopy screen.

7.7 Conclusions

I have been able to estimate the RMS accuracy of a 2D-3D registration algorithm, when registering clinical images which have no “gold-standard” registration, by using *k-fold* cross validation. My method for estimating RMS errors was tested using phantom data where it was shown to be a reasonable approximation to an RMS error calculated using a “gold-standard” registration. By registering to two fluoroscopy images I was able to remove the effects of spinal deformation. The results presented in this chapter back up my earlier findings from chapter 6 where phantom images were used. In chapter 6 the type ’2’ similarity measures were found to be susceptible to soft tissue structures. In this chapter the type ’2’ similarity measures were found to be the least accurate. In chapter 6 the pattern intensity and gradient difference measures were found to be the most accurate and robust. In this chapter all of the type ’1’ and ’3’ measures had fairly similar performance, though overall the pattern intensity and gradient difference measures appear slightly more accurate. Using the final registration results the amount of spinal deformation between the pre-operative and intra-operative images could be calculated. In a number of cases the amount of spinal deformation was much larger than the estimated registration accuracy. The largest rotational deformation was recorded in the flexion-extension direction and had a range of 5° over four

vertebrae.

Registrations were also carried out using a dataset from a percutaneous laser nucleotomy procedure. Although the CT scan had only a 13 slices, precise registrations which appeared visually accurate were achieved. A large local minima was found for 25% of the registrations. This is thought to be caused by a sagittal line of symmetry through the vertebra causing DRRs from two different angles to appear very similar. A registration between the pre-operative MR image and the intra-operative fluoroscopy image was achieved by initially registering the pre-operative MR image to the CT image and then registering the CT image to the fluoroscopy image. These registrations were used to project the nucleus pulpose and the prolapsed material from the MR onto the fluoroscopy image. The projected position of nucleus pulpose and prolapsed material appears visually correct.

Chapter 8

Deforming the pre-operative image to represent the inter-operative scene

8.1 Introduction

Soft tissue deformation can be a problem if a pre-operative modality is used to help guide a surgical or an interventional procedure. For example, consider the following situation. The 2D-3D registration algorithm has been used to register an individual vertebra in a pre-operative CT scan to the same vertebra in an interventional fluoroscopy image. The final set of extrinsic parameters are then used to project the position of a nerve, segmented from the CT volume, onto the fluoroscopy image. Assuming an accurate registration, and that the nerve and the vertebra have not moved relative to each other, then the projected position of the nerve should be accurate. However, if soft tissue deformation has occurred, which results in the nerve moving relative to the vertebra, then the rigid body assumption is no longer valid. This could cause an error in the projected position of the nerve. In this chapter I describe a method which can warp the pre-operative CT image to more accurately represent the intra-operative scene as shown by an interventional fluoroscopy image. This method has the potential to reduce the magnitude of any registration errors caused by soft tissue deformation because the rigid body assumption should be valid between the warped CT volume and the fluoroscopy image.

The previous chapter has shown that the 2D-3D registration algorithm can be used to calculate the amount of spinal deformation between a pre-operative CT scan and an intra-operative fluoroscopy image. The aim of this chapter is to use the information on the relative positions of different vertebrae (obtained using the 2D-3D registration algorithm) as an input into a deformation algorithm developed by Little *et al.* [75]. This deformation algorithm allows rigid bodies to

be incorporated into a non-linear deformation based on a radial basis function. The vertebrae are transformed as rigid bodies and the soft tissue structures surrounding the vertebrae are allowed to deform, based on the movement of the vertebrae and some additional landmark positions.

This chapter begins by describing the deformation algorithm by Little *et al.* [75]. This algorithm was found to produce unrealistic soft tissue deformations close to rigid structures. Section 8.2.5 describes an alteration to the deformation algorithm which improves the delineation of soft tissue structures close to the rigid bodies. In the second half of the chapter the deformation algorithm is used to warp a pre-operative CT volume so that it more accurately represents the intra-operative scene. Information on the amount of deformation between the pre-operative and intra-operative scenes is acquired by performing a number of 2D-3D registrations between a pre-operative CT volume and an intra-operative fluoroscopy image.

8.2 Deformation algorithm developed by Little *et al.* [75]

The algorithm described in Little *et al.* [75] allows rigid bodies to be incorporated in a non-linear deformation. The algorithm combines two interpolation techniques; point based interpolation and inverse distance weighted interpolation. Standard point based interpolation requires the whole image to be deformed as if it were a single homogeneous medium, while the algorithm described in Little *et al.* [75] allows some parts of the deformed image to be transformed solely using rigid transformations.

This section begins by describing point based interpolation. This is followed by a description of inverse distance weighted interpolation. These two interpolation techniques are combined to produce the deformation algorithm. The descriptions in this thesis have concentrated on the 3D case, for a more in-depth description see Little *et al.* [75], where the mathematics is described for an arbitrary number of dimensions. The final part of this section describes an improvement to the algorithm which produces more realistic soft tissue deformations close to rigid structures.

8.2.1 Point based interpolation

Point based interpolation is useful if there are a number of data values at given landmark locations and the task is to estimate data values at other locations. In three dimensions, given n data values $x_i, i = 1, \ldots, n$, at positions $\mathbf{w}_i = (w_i[1], w_i[2], w_i[3]), i = 1, \ldots, n$, then an interpolating function can be found which is exact at the landmark sites $f(\mathbf{w}_i) = x_i, i = 1, \ldots, n$ and is smooth and continuous elsewhere. Three methods which have been widely used lead to the same func-

tional form for the interpolator [95]. These are kriging, which provides the best unbiased predictor for a given stochastic process [83], thin-plate splines, which minimise a bending energy function [12] and radial basis functions, which provide an exact solution at the landmark sites, but have no minimisation constraints [91]. A general form of an interpolating function, for $\mathbb{R}^3 \rightarrow \mathbb{R}^1$, can be written as follows,

$$f(\mathbf{w}) = a_0 + \sum_{j=1}^{3} a_j w[j] + \sum_{i=1}^{n} b_i \sigma(\mathbf{w}, \mathbf{w}_i), \tag{8.1}$$

where the data value is estimated at position $\mathbf{w}$; a_0, a_1, a_2, a_3 are coefficients which represent the underlying linear transformation; $b_1, \ldots, b_n$ are coefficients which represent the non-linear component of the deformation and $\sigma(\mathbf{w}, \mathbf{w}_i)$ is a function which can take various forms depending on whether a kriging, thin-plate spline or radial basis function interpolator is used. An important special case is the thin-plate spline [12], where in 3D the basis function $\sigma(\mathbf{w}, \mathbf{w}_i)$ is given by $\sigma(\mathbf{w}, \mathbf{w}_i) = \|\mathbf{w} - \mathbf{w}_i\|$.

The coefficients $a_0, a_1, a_2, a_3, b_1, \ldots, b_n$ are determined using the following equation,

$$\begin{pmatrix} \Sigma & D \\ D^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix}, \tag{8.2}$$

where

$$\mathbf{x} = (x_1, \ldots, x_n)^T, \tag{8.3}$$

$$\mathbf{b} = (b_1, \ldots, b_n)^T, \quad \mathbf{a} = (a_0, a_1, a_2, a_3)^T, \tag{8.4}$$

$$D = \begin{pmatrix} 1 & w_1[1] & w_1[2] & w_1[3] \\ \vdots & \vdots & \vdots & \vdots \\ 1 & w_n[1] & w_n[2] & w_n[3] \end{pmatrix}, \tag{8.5}$$

$$\Sigma = \begin{pmatrix} \sigma(\mathbf{w}_1, \mathbf{w}_1) & \sigma(\mathbf{w}_1, \mathbf{w}_2) & \cdots & \sigma(\mathbf{w}_1, \mathbf{w}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(\mathbf{w}_n, \mathbf{w}_1) & \sigma(\mathbf{w}_n, \mathbf{w}_2) & \cdots & \sigma(\mathbf{w}_n, \mathbf{w}_n) \end{pmatrix} \tag{8.6}$$

The top row of the matrix equation (8.2), $\Sigma\mathbf{b} + D\mathbf{a} = \mathbf{x}$, ensures that the landmark constraints are satisfied, e.g. $f(\mathbf{w}_i) = x_i, i = 1, \ldots, n$. The bottom row of the matrix equation (8.2), $D^T\mathbf{b} = 0$, ensures that the deformation tends to zero as the distances between point $\mathbf{w}$ and the landmark positions, $\mathbf{w}_i, i = 1, \ldots, n$, tend to infinity.

This interpolation method can be applied to deform an image in the following manner. Starting with a set of initial landmarks, $\mathbf{w}_i = (w_i[1], w_i[2], w_i[3]), i = 1, \ldots, n$, which transform to new positions, $\mathbf{v}_i = (v_i[1], v_i[2], v_i[3]), i = 1, \ldots, n$, then three individual interpolators can be calculated, one for each of the three dimensions. The landmark constraints are $f_1(\mathbf{w}_i) = v_i[1], f_2(\mathbf{w}_i) = v_i[2], f_3(\mathbf{w}_i) = v_i[3], i = 1, \ldots, n$, and each interpolator has an individual set of coefficients, $a_0[l], a_1[l], a_2[l], a_3[l], b_1[l], \ldots, b_n[l]$ where $l = 1, 2, 3$. Therefore, an initial 3D point $\mathbf{w}$ will map to a final position $\mathbf{v}$ given by,

$$\mathbf{w} \rightarrow (f_1(\mathbf{w}), f_2(\mathbf{w}), f_3(\mathbf{w})) = \mathbf{v} \tag{8.7}$$

The interpolation functions are given by,

$$f_l(\mathbf{w}) = a_0[l] + \sum_{j=1}^{3} a_j[l] w[j] + \sum_{i=1}^{n} b_i[l] \sigma(\mathbf{w}, \mathbf{w}_i), \quad l = 1, 2, 3, \tag{8.8}$$

and the coefficients are given by the solution to the following equation,

$$\begin{pmatrix} \Sigma & D \\ D^T & 0 \end{pmatrix} \begin{pmatrix} B \\ A \end{pmatrix} = \begin{pmatrix} V \\ 0 \end{pmatrix}, \tag{8.9}$$

where the matrices D and Σ are the same as before, given by equations (8.5) and (8.6). The matrices A, B and V are the sets of interpolation coefficients, $\mathbf{a}$ and $\mathbf{b}$, and the set of landmark constraints, $\mathbf{v}$, respectively, see equation (8.10).

$$A = \begin{pmatrix} a_0[1] & a_0[2] & a_0[3] \\ \vdots & \vdots & \vdots \\ a_3[1] & a_3[2] & a_3[3] \end{pmatrix}, B = \begin{pmatrix} b_1[1] & b_1[2] & b_1[3] \\ \vdots & \vdots & \vdots \\ b_n[1] & b_n[2] & b_n[3] \end{pmatrix}, V = \begin{pmatrix} v_1[1] & v_1[2] & v_1[3] \\ \vdots & \vdots & \vdots \\ v_n[1] & v_n[2] & v_n[3] \end{pmatrix} \tag{8.10}$$

8.2.2 Inverse distance weighted interpolation

Inverse distance weighted interpolation methods were originally proposed by Shepard [104, 107]. Using the same notation as the previous section i.e. there are n data values at positions $\mathbf{w}_i = (w_i[1], w_i[2], w_i[3]), i = 1, \ldots, n$ with values $x_i, i = 1, \ldots, n$. To estimate a data value at position $\mathbf{w}$, a weighted average of the data values is calculated, where the weights λ_i, are dependent on the distance between the point $\mathbf{w}$ and each landmark position, $\mathbf{w}_i$. The interpolating function is given by,

$$f(\mathbf{w}) = \sum_{i=1}^{n} \lambda_i(\mathbf{w}) x_i, \tag{8.11}$$

and weighting function, λ, is chosen to have the following properties,

$$\begin{aligned} \lambda_i(\mathbf{w}_i) = 1, \quad & i = 1, \ldots, n \\ \sum_{i=1}^{n} \lambda_i(\mathbf{w}) &= 1 \\ \lambda_i(\mathbf{w}) \geq 0, \quad & i = 1, \ldots, n \end{aligned} \tag{8.12}$$

These properties ensure that the interpolation function is exact at the landmark sites and is a weighted average of the data values away from the landmark sites. The weighting function proposed by Shepard in [107] is,

$$\lambda_i(\mathbf{w}) = \frac{q_i(\mathbf{w})}{\sum_{i=1}^{n} q_i(\mathbf{w})} \qquad \text{where,} \qquad q_i(\mathbf{w}) = \frac{1}{\|\mathbf{w} - \mathbf{w}_i\|^{\mu}}, \tag{8.13}$$

and μ is a positive constant greater than one.

8.2.3 Deformations incorporating rigid structures

This section describes how the two different interpolation schemes are combined in Little *et al.* [75] to incorporate rigid structures into the deformation. The algorithm by Little *et al.* [75] ensures that the rigid bodies do not deform by using the following approach. Within a rigid body the underlying linear transformation is set to be the transformation of that rigid body and the non-linear component of the warp is set to equal zero. The underlying linear transformation part is controlled by replacing the first part of equation (8.8), $a_0[l] + \sum_{j=1}^{3} a_j[l]w[j]$, by the following inverse distance weighted interpolation scheme,

$$\sum_{i=1}^{n_o} \lambda_i(\mathbf{w}) L_i \mathbf{w}', \tag{8.14}$$

where there are n_o rigid bodies which each have a transformation matrix, L_i and $\mathbf{w}'$ is the point $\mathbf{w}$ in homogeneous coordinates e.g. $\mathbf{w}' = (w[1], w[2], w[3], 1)^T$. The weighting factor λ is calculated as shown by equation (8.13), except that the distance between points, $\|\mathbf{w} - \mathbf{w}_i\|$, is replaced by $\mathcal{S}_i(\mathbf{w})$, which represents the minimum distance between point $\mathbf{w}$ and the i'th rigid body. The positive constant, μ, determines the smoothness of the interpolation. The value μ = 2 is a popular choice [104], though μ = 1.5 was used for the experiments described in this thesis to make the algorithm consistent with the work described in Little *et al.* [75].

The non-linear part of the deformation is constrained to equal zero within a rigid body by ensuring that the basis function, $\sigma(\mathbf{w}, \mathbf{w}_i)$, equals zero within a rigid body. This is achieved by multiplying the basis function by, $\mathcal{S}_0(\mathbf{w})\mathcal{S}_0(\mathbf{w}_i)$, see equation 8.15.

$$\sigma'(\mathbf{w}, \mathbf{w}_i) = \mathcal{S}_0(\mathbf{w})\mathcal{S}_0(\mathbf{w}_i)\sigma(\mathbf{w}, \mathbf{w}_i), \tag{8.15}$$

where $\mathcal{S}_0(\mathbf{w})$ represents the minimum distance between the point $\mathbf{w}$ and the nearest rigid body. Therefore, the interpolating solution for deformations incorporating rigid structures is given by,

$$f(\mathbf{w}) = \sum_{i=1}^{n_o} \lambda_i(\mathbf{w}) L_i \mathbf{w}' + \sum_{j=1}^{n} B_j^T \sigma'(\mathbf{w}, \mathbf{w}_j), \tag{8.16}$$

where the first term is the inverse distance weighted interpolation scheme and the second term corresponds to the point based deformation where B_j represents the j'th row of matrix B. The coefficients of the matrix B are found using the following equation.

$$B = \Sigma^{-1}(V - T), \tag{8.17}$$

where,

$$T = \begin{pmatrix} (\sum_{i=1}^{n_o} \lambda_i(\mathbf{w}_1) L_i \mathbf{w}_1')^T \\ \vdots \\ (\sum_{i=1}^{n_o} \lambda_i(\mathbf{w}_n) L_i \mathbf{w}_n')^T \end{pmatrix} \tag{8.18}$$

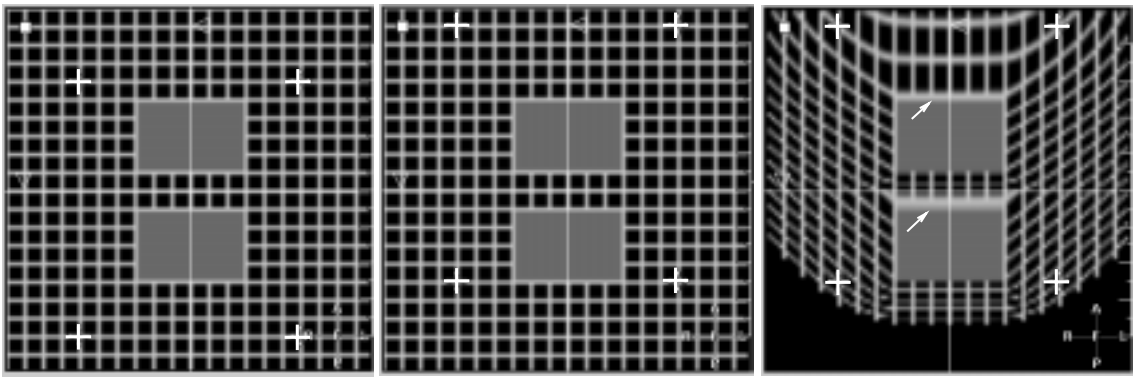
The matrix T is necessary to ensure that the landmark constraints are satisfied. The radial basis function σ', used in the deformation scheme (equation (8.16)) is a modified Hardy multiquadric, given by,

$$\sigma'(\mathbf{w}, \mathbf{w}_i) = \mathcal{S}_0(\mathbf{w})\mathcal{S}_0(\mathbf{w}_i)\sqrt{(\|\mathbf{w} - \mathbf{w}_i\|^2 + \mathcal{S}_0^2(\mathbf{w}_i))} \tag{8.19}$$

8.2.4 Unrealistic soft tissue deformations close to rigid bodies

The original implementation of the deformation algorithm by Little *et al.* [75] was found to produce unrealistic deformations in soft tissue structures close to rigid bodies. An example, using a computer generated image, which contains a grid pattern and two rectangular rigid bodies, is shown in figure 8.1. The source image (figure 8.1(a)) has been warped to the target image (figure 8.1(b)) and the resultant image is shown in figure 8.1(c). The warp is based on four landmarks, which move vertically upwards between the source image and the target image, and two rigid

bodies which remain stationary. Figure 8.1(c) shows a smooth deformation outside of the rigid bodies; zero deformation within the rigid bodies and the landmarks have moved exactly. However, there are large deformations close to the rigid bodies. In particular, there are large deformations immediately above the rigid bodies and the midline between the two rigid bodies has moved upwards. If the rigid bodies in figure 8.1 were two vertebrae and the space between them was occupied by an intervertebral disc, I would expect the deformation of the disc to be primarily influenced by the movement of the surrounding vertebrae. However, figure 8.1(c) shows a large amount of deformation between the two rigid bodies, even though the rigid bodies themselves have not moved.



(a) (b) (c)

Figure 8.1: Illustrating an unrealistic deformation close to the rigid bodies. The source image (a) has been warped to the target image (b) based on the four landmarks shown by white crosses and two rigid bodies shown by the grey rectangles. The rigid bodies do not move, the landmarks all move vertically upwards. The warped image (c) shows large deformations close to the rigid bodies. These are particularly evident immediately above the rigid bodies, marked by the arrows in (c).

A demonstration of how the algorithm by Little *et al.* [75] can cause unrealistic deformations in intervertebral discs is shown in figure 8.2. This figure shows the results of an experiment to warp an MR volume of a volunteer with their neck in a flexed position, figure 8.2(a), into the extended position shown by figure 8.2(b). The seven cervical vertebrae were used as rigid bodies. To constrain the point based part of the algorithm 22 corresponding anatomical point pairs were identified by a radiologist in both the flexed and extended images. The deformed image, figure 8.2(c), was examined by a radiologist to ascertain how realistically the soft tissue structures had deformed. Unrealistic deformations were observed in the soft tissue structures close to the rigid bodies. This was particularly apparent when the positions of the intervertebral discs were examined. The discs often appeared to be squashed against one of the neighbouring vertebrae and in some cases had virtually disappeared.

(a) (b) (c)

Figure 8.2: Magnified sagittal sections through MR volumes of the source image (a), the target image (b) and the warped image (c). The deformation was based on 22 manually picked landmarks and seven rigid bodies, the cervical vertebrae. Note that the intervertebral discs in the warped image are blurred and in some cases have almost disappeared, while the discs are clearly visible in the source (a) and target (b) images.

8.2.5 Improving the delineation of soft tissue structures close to rigid bodies

This section describes how the algorithm was modified to improve the delineation of soft tissue structures close to rigid bodies. The radial basis function was changed to a standard Hardy multiquadric, $\sigma(\mathbf{w}, \mathbf{w}_j) = \sqrt{\|\mathbf{w} - \mathbf{w}_j\|^2 + \mathcal{S}_0^2(\mathbf{w}_j)}$ and a weighting parameter Ω, was introduced to control the weighting between the two deformation schemes. The interpolating function, which was previously given by equation (8.16), is now given by,

$$f(\mathbf{w}) = \sum_{i=1}^{n_o} \lambda_i(\mathbf{w}) L_i \mathbf{w}' + \Omega(\mathbf{w}) \sum_{j=1}^{n} B_j^T \sigma(\mathbf{w}, \mathbf{w}_j) \tag{8.20}$$

and the weighting factor is given by,

$$\Omega(\mathbf{w}) = \frac{\mathcal{S}_0^{\alpha}(\mathbf{w})}{\mathcal{S}_0^{\alpha}(\mathbf{w}) + \beta|\mathbf{w} - \mathbf{w}_{nearest}|} \quad \text{outside a rigid body} \qquad \Omega(\mathbf{w}) = 0 \quad \text{inside a rigid body,} \tag{8.21}$$

where $\mathbf{w}_{nearest}$ represents the position of the landmark which has the smallest Euclidean distance to point $\mathbf{w}$ and the exponent α and the scalar β are positive constants. The weighting function $\Omega(\mathbf{w})$ has been chosen so that away from the rigid bodies the non-linear deformation dominates

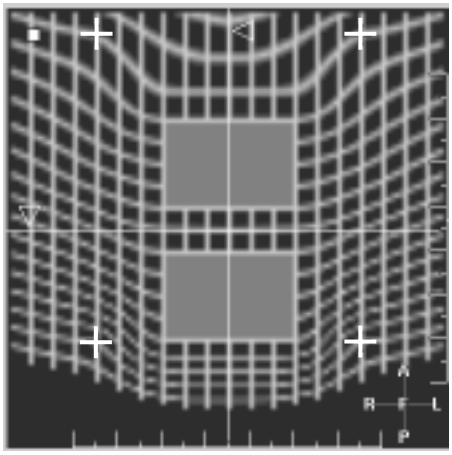
(i.e. for large $\mathcal{S}_0$, $\Omega(\mathbf{w}) \rightarrow 1$) and also so that, as a point approaches a rigid body the weighting factor tends to zero and the inverse distance weighted interpolation scheme dominates (i.e. as $\mathcal{S}_0 \rightarrow 0$, $\Omega(\mathbf{w}) \rightarrow 0$). At a landmark position $\Omega(\mathbf{w}) = 1$ and so the landmarks are transformed exactly. At a rigid body $\Omega(\mathbf{w}) = 0$ and so the transformation is governed by the inverse distance weighted interpolation scheme, which is a rigid transformation within a rigid body. The values of α and β determine how rapidly the interpolation scheme alters between the inverse distance weighted interpolation close to a rigid body and the point based interpolation at a distance from the rigid bodies.

The original implementation of the deformation algorithm controlled the weighting between the two deformation schemes by using the modified Hardy multiquadric shown in equation (8.15). Although this implementation satisfied the mathematical constraints (exact at landmark sites and non-linear deformation tends to zero at rigid bodies) there are a number of equations which can also achieve these constraints, some of which will produce more realistic soft tissue deformations than others. By explicitly defining the weighting function between the deformation schemes I was able to tailor the weighting function to produce visually realistic deformations [97]. The values $\alpha = 3$ and $\beta = 10$ were used for the experiments described in this thesis.

The advantage of increasing the number of adjustable parameters into the deformation algorithm is the ability to tailor the algorithm to obtain visually realistic deformations close to the rigid bodies. The dangers of increasing the number of adjustable parameters is that the deformation algorithm may become tailored to one specific set of data and will produce erroneous results when applied to different data. The values of α and β were obtained by visual inspection of the registration between the flexed and extended MR image volumes of the neck shown in figure 8.2 [97]. These values of α and β have subsequently been applied to three other pairs of MR images of the neck, one pair of MR knee images and, as described in the following section, to deform a CT volume of the lumbar spine. In each case the position of soft tissue structures close to the rigid bodies appeared visually realistic.

A second alteration to the algorithm was also found to make the deformations close to rigid bodies more realistic. In the original implementation the rigid bodies had very little influence on the landmark based deformation. To correct for this, each rigid body is represented in the landmark based deformation scheme by adding an additional landmark, which is found automatically at the centre of gravity of the rigid body. This landmark was transformed between the target and source images using the transformation matrix for that particular rigid body. In the original implementation of the algorithm it was not possible to include landmarks on rigid bodies as this would

have resulted in the matrix Σ (see equation (8.17)) being singular and hence non-invertible. In the new implementation the radial basis function does not equal zero when $\mathcal{S}_0(\mathbf{w}) = 0$ and so the matrix Σ can still be inverted when landmarks are placed on the rigid bodies.



(a)

(b)

Figure 8.3: Deformation results using the improved algorithm. Results from deforming the image of a grid with two rigid bodies is shown in (a). Note that deformations close to the rigid bodies are now much smaller. Results from deforming the MR image of the volunteer is shown in (b). Note that the intervertebral discs are now much clearer, compared to figure 8.2(c).

The effect of these modifications are shown in figure 8.3. Figure 8.3(a) shows the result of warping the image of a grid which contains two rigid bodies. Again the landmarks are moved exactly and there is no deformation within the rigid bodies, however, now there is much less deformation close to the rigid bodies. In particular there is very little deformation between the two rigid bodies. By altering the weighting between the two deformation schemes and by including landmarks on the rigid bodies the influence of the rigid bodies on the surrounding soft tissue structures has been increased. This can produce more realistic deformations close to the rigid bodies, for example, figure 8.3(b) shows the result when warping the volunteer MR image. The intervertebral discs in figure 8.3(b) are much clearer than in figure 8.2(c).

8.3 An experiment to deform a pre-operative CT volume to more accurately represent the intra-operative scene

This experiment uses the 2D-3D registration algorithm to obtain information on the amount of spinal deformation that has occurred between a pre-operative CT volume an interventional fluoroscopy image. This information is used to deform the pre-operative image so that it more accurately represents the intra-operative scene. The CT volume from dataset 3 (see section 7.5.4) was used as the pre-operative image and the fluoroscopy image shown in figure 8.4 was used to represent the intra-operative scene. Dataset 3 was chosen because it had the largest amount of spinal deformation, as shown by the registration results in sections 7.5.2 to 7.5.5.

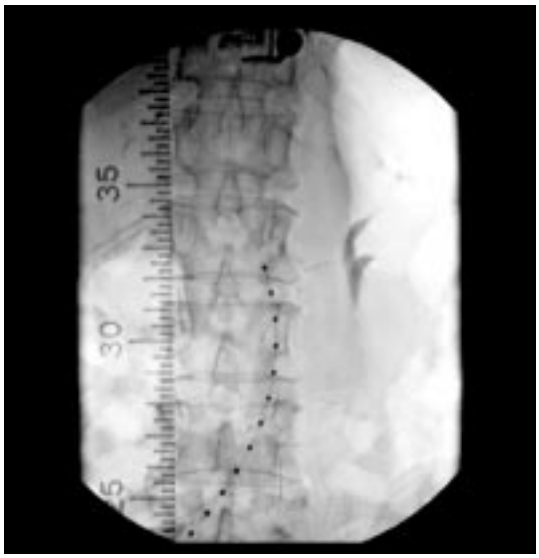


Figure 8.4: Distortion corrected fluoroscopy image. This image represents the intra-operative scene. It is used, along with the 2D-3D registration algorithm and the pre-operative CT volume, to provide information on the amount of spinal deformation which has occurred between the pre-operative image and the intra-operative scene.

8.3.1 Constraining the inverse distance weighted interpolation

The inverse distance weighted interpolation part of the algorithm is constrained by defining a number of rigid bodies which have associated transformation matrices, L_i. Four lumbar vertebrae were used as rigid bodies, L1, L2, L3 and L4. These vertebrae were segmented from the CT volume using the Philips EasyVision segmentation tool. The method involved manually drawing around the edge of the vertebra in each slice of the CT volume. Slices through the segmented image can be seen in figure 8.5.

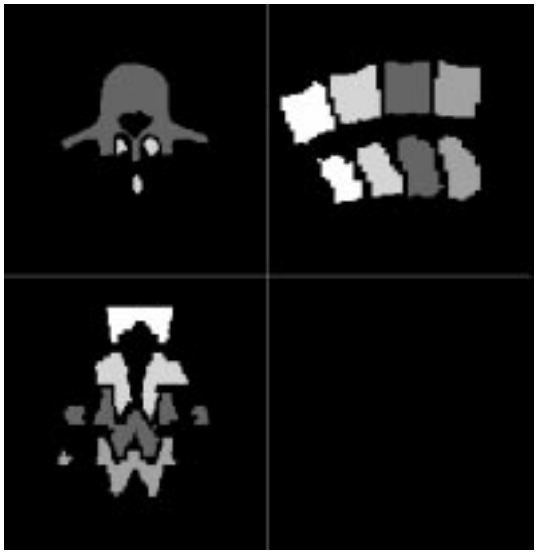


Figure 8.5: Axial, sagittal and coronal sections through a segmented CT image which contains four lumbar vertebrae, L1 $\rightarrow$ L4.

The 2D-3D registration algorithm was used to obtain the transformation matrices L_i. The method was as follows. A reduced CT volume was drawn around each vertebra and segmented from the CT volume, and a region of interest was manually defined around each vertebra in the fluoroscopy image. A starting estimate was obtained using the method described in section 7.5.1 and a single registration was carried out to each vertebra using the gradient difference similarity measure. The four sets of final extrinsic parameters were translated to represent a rotation around the centre of the whole CT volume rather than around a reduced CT volume. This gave four sets of extrinsic parameters, $\mathbf{E}_i, i = 1, \ldots, 4$, which, assuming that all four registrations are accurate, would be identical if there was no spinal deformation. A mean set of parameters were calculated, $\overline{E^k} = \frac{1}{4}\sum_i^4 E_i^k, k = 1, \ldots, 6$. The deviations from the mean set of parameters were also calculated for each vertebrae, $\delta E_i^k = \overline{E^k} - E_i^k, i = 1, \ldots, 4$ and $k = 1, \ldots, 6$. These deviations were used to construct the rigid body transformation matrices for each vertebrae, $L_i, i = 1, \ldots, 4$.

8.3.2 Constraining the landmark based deformation

Landmarks were automatically chosen at the centre of gravity of each rigid body. These points are transformed between the pre-operative and the intra-operative positions using the rigid body transformation matrices, $L_i, i = 1, \ldots, 4$. The intra-operative fluoroscopy image, see figure 8.4, shows very little soft tissue detail and so it was not possible to pick any landmarks to represent how soft tissue structures deform between the pre-operative and intra-operative scene.

Interpolation schemes can extrapolate information poorly and so can produce very unrealistic deformations outside the convex hull of the data points. Because of this, eight landmarks were picked to anchor the corners of the CT volume, as shown by the white crosses in figure 8.6. The assumption is, that the movement of the rigid bodies represented using both the inverse distance weighted interpolation and the landmarks in the rigid bodies, will result in a realistic deformation for the soft tissue structures in close proximity to the rigid bodies.

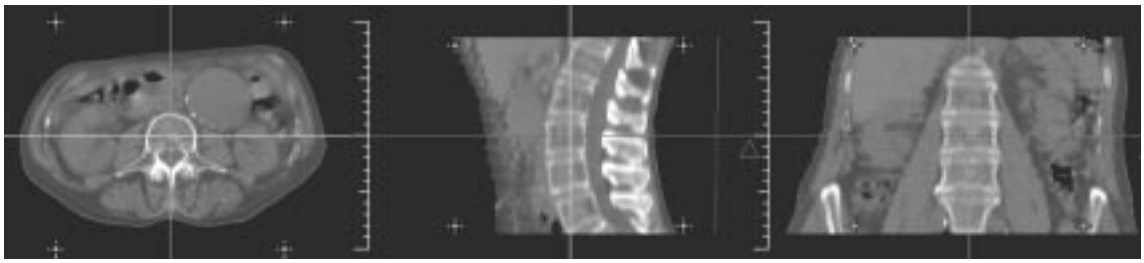


Figure 8.6: Axial, sagittal and coronal sections through pre-operative CT. White crosses show where the eight landmarks were chosen on the source and target images. These landmarks were chosen to increase the size of the convex hull of the data points to prevent the interpolation scheme producing "wild" deformations within the CT volume.

8.4 Results

Table 8.1 shows the mean set of final extrinsic parameters, $\overline{\mathbf{E}}$, and also the deviation from $\overline{\mathbf{E}}$ required to obtain the final set of parameters for each vertebrae, $\delta\mathbf{E}_i, i = 1, \ldots, 4$. These deviations have been used to calculate the rigid body transformation matrices, $L_i, i = 1, \ldots, 4$. However, because the algorithm is much less accurate in the out-of-plane translation direction X, the deviations from the mean for the X translation parameter were set to zero when the transformation matrices, $L_i, i = 1, \ldots, 4$, were calculated.

Vertebra	extrinsic parameters					
	rotation (degrees)			translation (mm)		
	θ_x	θ_y	θ_z	X	Y	Z
$\overline{\mathbf{E}}$	0.87	-0.21	-94.09	869.6	35.15	21.65
$\delta\mathbf{E}_1$ (L1)	0.03	2.28	0.78	1.7	-0.37	-2.35
$\delta\mathbf{E}_2$ (L2)	0.09	0.83	-0.20	3.7	0.33	-1.16
$\delta\mathbf{E}_3$ (L3)	0.01	-0.11	-0.07	-2.5	-0.22	0.89
$\delta\mathbf{E}_4$ (L4)	-0.13	-3.00	-0.50	-2.8	0.26	2.63

Table 8.1: The average set of final extrinsic parameters $\overline{\mathbf{E}}$, and the deviation from this average required to obtain the final set of extrinsic parameters to each individual vertebra $\delta\mathbf{E}_i, i = 1, \ldots, 4$.

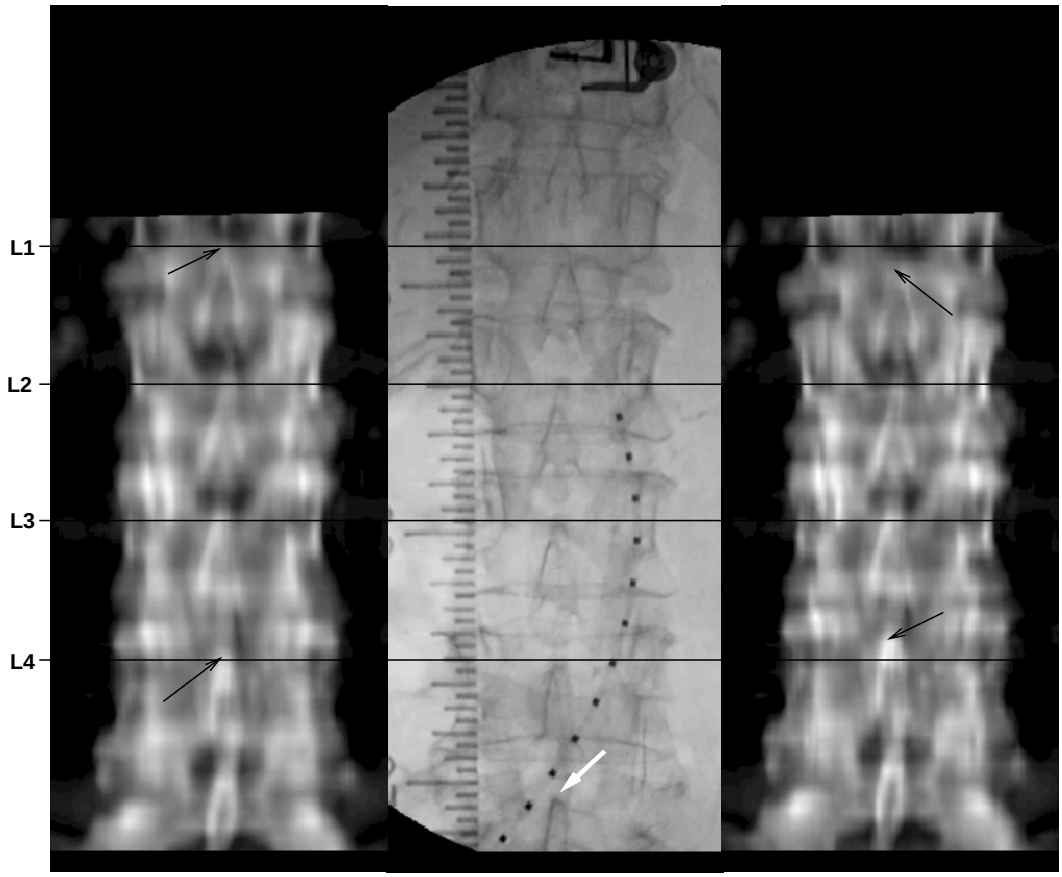
Axial and sagittal sections through the warped and original CT volumes can be seen in figure 8.7. The major deformation is a flexion-extension of the spine, where the pre-operative CT volume shows the spine in a more extended position than the warped CT volume. The vertebrae have transformed as rigid bodies and the adjacent soft tissue structures appear to have deformed in a visually realistic manner.

(a) (b)

Figure 8.7: Axial and sagittal sections through the deformed CT volume (a) and through the pre-operative CT volume (b). The L1, L2, L3 and L4 vertebrae are shown. Note that the vertebrae have moved as rigid bodies. The pre-operative CT volume shows more extension in the spine than the deformed volume.

DRRs were produced from the pre-operative CT volume and from the warped CT volume using the average final extrinsic parameters, $\overline{\mathbf{E}}$. By visual inspection (overlaying each DRR in turn onto the fluoroscopy image) the DRR produced using the warped CT volume appeared to be more accurately registered to the fluoroscopy image than the DRR produced using the original CT volume. This was particularly evident if the L1 and L4 vertebrae were examined. It is difficult to show these overlays in this thesis because a major tool in carrying out the visual inspection was the ability to fade one of the overlays in and out of the overall image. Instead, the DRRs are shown alongside the interventional fluoroscopy image in figure 8.8. To compare the three images, four horizontal lines have been drawn, which indicate the vertical position of the superior surface of the L1, L2, L3 and L4 spinous processes in the fluoroscopy image (this feature is indicated by a white

arrow on the L5 vertebra). The vertical positions of the superior surfaces of the spinous processes in the two DRRs were compared with these horizontal lines. The superior surfaces of the spinous processes in the DRR produced using the warped CT volume, as shown in figure 8.8(a), can be seen to lie closer to their respective black lines than the superior surfaces of the spinous processes in the DRR produced using the original CT volume, shown in 8.8(c). This is particularly evident for the L1 and L4 spinous processes, which have been indicated by the black arrows in figure 8.8. This shows that the DRR produced using the warped CT volume is more accurately registered to the fluoroscopy image than the DRR produced using the original CT volume.



(a) (b) (c)

Figure 8.8: DRR produced with warped CT volume (a), intra-operative fluoroscopy image on which warp was based (b) and DRR produced with pre-operative CT volume (c). DRRs were produced using an average set of extrinsic parameters from registrations to the L1, L2, L3 and L4 vertebrae. The black lines represent the vertical position of the superior surface of the spinous processes in the fluoroscopy image.

8.5 Conclusions

Deformation between the pre-operative and the intra-operative scene can be a problem if pre-operative modalities are used to help guide interventional procedures. This chapter has described a novel combination of a 2D-3D registration algorithm and a deformation algorithm. The 2D-3D registration algorithm is used to provide information to the deformation algorithm so the pre-operative image can be deformed to more accurately represent the intra-operative scene. A DRR produced using the deformed CT volume visually appeared to be more accurately registered to the fluoroscopy image than a DRR produced using the original CT volume. However, only the position of bony structures could be compared. Very little can be said about the registration accuracy of soft tissue structures, because very little information on the intra-operative position of soft tissue structures is available from the fluoroscopy image. However, in a previous experiment using the deformation algorithm, when both the target and source images were 3D MR scans [97], the algorithm produced visually realistic soft tissue deformations close to the rigid bodies.

This chapter has described preliminary work which has undergone relatively little validation, especially when it is compared with the extensive validation experiments described in the previous two chapters. Improved validation strategies for the deformation algorithm are discussed in the section on future work, section 9.5.1.

Chapter 9

Conclusions and future work

The main objective of the work described in this thesis was to design a 2D-3D image registration algorithm to register a pre-operative CT image to an interventional fluoroscopy image. The clinical motivation behind this work was to improve image guidance during interventional procedures. This chapter will summarise how close I have come to achieving this objective, will highlight the novel contributions of the work and will suggest future work.

9.1 The registration algorithm

As a result of the literature review (chapter 3) a decision was made to design an intensity based algorithm. Chapter 4 describes how DRRs are formed from the CT volume and how the fluoroscopy image is altered so that it more accurately resembles the DRR. The novel contributions of my approach are as follows. Only voxels containing bone are used to produce the DRR. Therefore, the registration is based on bony detail and so soft tissue deformation is not an issue. A bilinear interpolation method is used to estimate Hounsfield values within the CT volume. This method produces approximately the same results as trilinear interpolation but has a lower computational cost. A reduced CT volume was segmented from the CT volume and a ROI was defined within the fluoroscopy image. These were used to speed up the algorithm and reduce the number of truncated rays. The ROI in the fluoroscopy image can also be used to concentrate the algorithm on a region where the most accurate registration is required.

Chapter 5 describes the similarity measures, search space and search strategy used by the algorithm. Seven similarity measures were described, one of which (local correlation) has not been previously used for 2D-3D registration and one new measure, gradient difference. The design of the gradient difference measure has led to an extension of a patent application filed for the pattern

intensity similarity measure [132].

One of the breakthroughs in my work on similarity measures occurred when I began to investigate the ability of similarity measures to cope with the presence of differences between DRRs and fluoroscopy images. Previously I had concentrated on the ability of similarity measures to use common information between the two images. This is a key point because, by concentrating on the differences rather than the similarities between the images, I found it much easier to determine why some measures were more appropriate than others for registering CT volumes to interventional fluoroscopy images.

An experiment in chapter 5 showed that altering some of the intrinsic parameters can have nearly exactly the same effect on the appearance of the DRR as altering some of the extrinsic parameters. Because of this the search space chosen for the algorithm was the defined by the six extrinsic parameters and a full ten degree of freedom search was not attempted.

Chapter 5 also highlighted the importance of decoupling the translational parameters. An investigation was carried out which showed that fewer local minima were formed if the out-of-plane translation direction was along an axis from the x-ray source to the centre of the CT volume, as opposed to in a direction which is parallel to the X-axis. This resulted in the registration algorithm having non-orthogonal translational axes.

Table 9.1 provides a summary of the registration algorithm described in this thesis. This can be compared to the summary tables 3.1 and 3.2 in the review chapter on 2D-3D registration algorithms.

9.2 Comparison of similarity measures

One of the major contributions to knowledge described in this thesis is the comparison of different similarity measures. Experiments were carried out to determine how the accuracy and robustness of the registration algorithm was effected by using different similarity measures. These investigations were carried out using images of a lumbar spine phantom to which clinical features were added (in chapter 6) and using routinely acquired clinical images from aortic stenting procedures (in chapter 7).

9.2.1 Comparison of similarity measures using phantom images

Phantom images were initially used because it is much easier to obtain an accurate “gold-standard” registration using a phantom, compared to using clinical data. The “gold-standard” registration

Algorithm authors	Penney *et al.* 2000
Information extracted	Fluoroscopy images and DRRs
Similarity metric	Pattern intensity and Gradient difference
Search space	extrinsic
Search strategy	Gradient descent
Speed of registration	74 secs
Computer Used	SUN Ultra 30 (300MHz)
Images used	1. Fluoroscopy images and CT volume of lumbar spine phantom. 2. Routinely acquired clinical CT and fluoroscopy images.
Segmentation Method	Threshold in CT. ROI manually defined in CT and fluoroscopy images.
Number of views	1
Method of determining accuracy	Fiducial markers and *k-fold* cross validation.
Error measured in 2D or 3D	2D and 3D
Registration Accuracy	$\overline{d(ROI)} \leq$ 0.63mm 2D error using phantom images. Rotational errors and in-plane translational errors $\leq 1^{\circ}$ or mm and out-of-plane translational error $\leq$ 10mm using clinical images

Table 9.1: Summary of the 2D-3D registration algorithm described in this thesis.

was calculated using fiducial markers which were placed on the phantom prior to image acquisition. The "gold-standard" registration accuracy was estimated to be 0.09° for the rotational parameters, 0.15mm for the in-plane translational parameters and 1.34mm for the out-of-plane translational parameter.

Some of the problems involved with 2D-3D image registration for use in image guided interventions were not modelled by the images of the spine phantom alone. Because of this, more realistic images were simulated by overlaying features from clinical fluoroscopy images on to the fluoroscopy image of the spine phantom. The starting positions for the registration algorithm were chosen to simulate the expected initial error when registering in a clinical situation. Registration accuracy was measured in two ways, firstly as a RMS error in each of the extrinsic parameters and secondly as an overall error measure, reprojection distance. This was calculated over two areas, firstly over the entire fluoroscopy image, $\overline{d}$, and also over the fluoroscopy ROI, $\overline{d(ROI)}$.

All the similarity measures produced reasonably accurate registrations to the plain fluoroscopy image. The maximum value of $\overline{d(ROI)}$ was 0.96mm which is acceptable for all the interventional procedures described in chapter 2. The mutual information measure had a fairly high

failure rate, 25%, though the other measures were all robust, the maximum failure rate was 5.1% and a number of similarity measures recorded no failures. The most accurate and robust measures were those which either used gradient information or a local region. Mutual information was by far the worst measure in terms of accuracy and robustness. Mutual information is thought to be failing as there are far fewer pixels involved in our 2D-3D registration compared to the number of voxels in a typical 3D-3D registration scenario. This leaves the 2D histogram relatively sparsely populated, compared to the 1D histogram used by entropy of the difference image.

The introduction of soft tissue structures and interventional instruments was shown to have a large effect on the performance of some similarity measures which have been previously applied to 2D-3D registration. Registrations to the simulated image with soft tissue structures overlaid showed that measures which did not use gradient information or a local region had a high failure rate and large registration errors. The presence of the interventional stent was seen to effect the performance of measures where the maximum contribution from a given pair of pixels is strongly dependent on the pixel intensity (e.g. cross correlation and gradient correlation). Overall the two most accurate similarity measures were pattern intensity and gradient difference. They were both able to achieve accurate registrations ($\overline{d(ROI)}$ equals 0.59mm for pattern intensity and 0.63mm for gradient difference) and were reasonably robust (a 5.9% failure rate for pattern intensity and 10.2% for gradient difference) when registering to the fluoroscopy image containing both soft tissue structures and a stent.

9.2.2 Comparison of similarity measures using clinical images

In chapter 7 the comparison of similarity measures was continued using clinical images from aortic stenting procedures. However, this caused additional problems, the two most important of which were the lack of a "gold-standard" registration and the presence of spinal deformation. My solution to the lack of a "gold-standard" registration was to use *k-fold* cross validation to estimate RMS registration errors. To account for spinal deformation the method uses two fluoroscopy images, taken from different positions. A phantom experiment was carried out to calculate how accurately this approach could estimate RMS registration errors. The results of this experiment were that errors calculated using the *k-fold* cross validation approach were a reasonable approximation to RMS errors calculated using an independent "gold-standard" (i.e. the fractional error was less than 25%).

Experiments were carried out using four datasets. The results showed that the three measures which did not use gradient information or a local region were consistently the least accurate of

the seven measures. The four measures which used gradient information or a local region all produced accurate registrations (overall RMS rotational error of 0.76° and RMS in plane translational error of 0.85mm). These measures were also fairly robust, failure rate around 10%. The local correlation and gradient correlation measures performed better than had been expected from my earlier investigations using phantom images. However, the same two measures, pattern intensity and gradient difference, still appear to be the most accurate.

During these experiments the amount of spinal deformation was also calculated and, for some degrees of freedom, was observed to be significantly larger than the size of the registration errors. This shows that it was necessary to use the two view approach to remove the effects of spinal deformation.

The final section of chapter 7 registers images from a percutaneous laser nucleotomy procedure. Only the gradient difference and pattern intensity measures were used. These measures were able to register with high precision despite there being only 13 slices of CT data available. DRRs produced at the mean final registration position appeared, using visual inspection, to be accurately aligned with the fluoroscopy image. However, a problem was found in the form of a local minima which is believed to be caused by a sagittal line of symmetry through the vertebra. The patient also had a pre-operative MR scan which was registered to the pre-operative CT volume using a 3D-3D registration algorithm based on normalised mutual information [112, 116]. This enabled information, in the form of the nucleus pulpose and prolapse, to be overlayed from the MR image onto the fluoroscopy image.

9.3 Deformation

In chapter 8 the 2D-3D registration algorithm was used to acquire information on the amount of spinal deformation between a pre-operative CT image and an intra-operative fluoroscopy image. This information was used as an input into a deformation algorithm developed by Little *et al.* [75], which warped the pre-operative CT volume so that it more accurately represented the intra-operative scene.

The warped CT volume appeared visually realistic. To assess the accuracy of the deformed CT image DRRs were produced using both the original and the warped CT volumes. These DRRs were compared to the interventional fluoroscopy image. All of the vertebrae in the DRR produced using the warped CT volume visually appeared to be accurately registered to the fluoroscopy image, while errors were observed in the registration between some of the vertebrae in the DRR

produced using the original CT volume and the vertebrae in the fluoroscopy image. It is difficult to assess the registration accuracy of the soft tissue structures as very little information on the intra-operative position of these structures was available. However, the position of soft-tissue structures within the warped CT volume looked visually realistic.

9.4 Clinical impact

The technical side of the work carried out in this thesis has been very successful. The algorithm is able to accurately register clinical CT and fluoroscopy images. Registration errors are expected to be smaller than 1mm over a region of interest. Registrations can be achieved in a few minutes and require no prior segmentation apart from defining a rough ROI.

The algorithm has not yet, however, been able to effectively show its worth in a clinical setting within our hospital. From the three interventional procedures described in chapter 2, I believe that the treatment of abdominal aortic aneurysms is the procedure which is least likely to benefit from the use of a 2D-3D registration algorithm. This is because the stent is constrained within the aorta and so the positioning of the stent is essentially a one dimensional problem. The other two procedures, percutaneous ablation of liver metastases and the percutaneous laser nucleotomy, involve more 3D localisation and so are more likely to benefit from the use of a 2D-3D registration algorithm. At Guy's Hospital the percutaneous treatment of liver metastases is currently carried out using CT guidance and so no fluoroscopy images were available to demonstrate the potential of the algorithm. Although the percutaneous laser nucleotomy procedures are carried out using a combination of CT and fluoroscopy guidance, because these procedures have currently been put on hold due to poor clinical results I have not been able to carry out registrations during a procedure.

Other possible clinical benefits of the work described in this thesis are the use of the algorithm in other clinical areas such as neurosurgery, spinal surgery and hip-replacement. The voxel based nature of the algorithm means that it should be able to register images from any region of the body. It would also be interesting to apply the pattern intensity and gradient difference measures to radiotherapy images. I am unsure how these measures would perform when attempting to register CT volumes to portal images, because portal images have a very different contrast to fluoroscopy images. However, the problems involved in registering CT volumes to simulator images are very similar to registering CT volumes to interventional fluoroscopy images. Simulator images contain soft tissue structures and also often contain thin line structures such as delineator wires. Both the pattern intensity and gradient difference similarity measures are able to register accurately despite

the presence of soft tissue and thin line structures. Therefore, these measures may have great potential to aid radiotherapy patient positioning.

9.5 Future work

This section is split into two parts; future work concerning the validation of the deformation algorithm and future work required to use the algorithm in a clinical setting.

9.5.1 Validation of the deformation algorithm

The combination of the 2D-3D registration algorithm and the deformation algorithm was validated using visual assessment on the position of bony structures. It was difficult to assess the registration accuracy of soft tissue structures as very little information on the intra-operative position of these structures was available from the fluoroscopy image. A validation procedure for the position of soft tissue would require an intra-operative modality which can observe soft tissue structures. For example, images could be used from a procedure which acquires a pre-operative CT volume and is carried out using a combination of CT and fluoroscopy guidance. The pre-operative CT could be warped, based on the position of the vertebrae in the fluoroscopy image, and then the intra-operative CT volume could be used for validation purposes. The position of structures in the intra-operative CT volume and in the warped CT volume could be compared, either by using visual comparison, or a more quantitative measure could be used, such as the mean distance between structures of interest in both images.

9.5.2 Use of the algorithm in a clinical setting

Registration accuracy using two views

The work described in this thesis has concentrated on obtaining an accurate registration between a single fluoroscopy image and a CT volume. One of the major uses of this algorithm would be to localise a 3D point within the CT volume. This would require two 2D-3D registrations to fluoroscopy images taken from different views and then the same 2D point (such as the tip of a needle) would need to be localised in both fluoroscopy images. However, further work is required to ascertain the accuracy of this 3D localisation problem. In particular, how does the magnitude of the rotation (and/or translation) between the two fluoroscopy views relate to how accurately a 3D point can be localised within the CT volume? There are also issues involved with how accurately

the needle tip needs to be located in the fluoroscopy images and whether this accuracy can be achieved manually or is a feature extraction process required?

Detecting misregistrations and tracking

If the 2D-3D registration algorithm is used for clinical guidance then a system should be in place to detect any misregistrations. Misregistrations can occur due to a number of reasons e.g. if the starting position for the algorithm is outside of the capture range, or if a local minimum is present. Also, large mistakes can occur when defining the starting position, such as registering to the wrong vertebra or mistaking an AP view for a PA view.

An obvious system for detecting misregistrations is visual screening, where the interventionist would be required to visually compare the final DRR with the fluoroscopy image and to reject the registration if they are concerned with its accuracy. In most cases the misregistration can be visually observed. The production of a difference image as shown in figure 5.1 can help to visually highlight a misregistration. However, some misregistrations can be very difficult to detect visually.

Another possible method to detect misregistrations is to set a threshold for the final value of the similarity measure and if this threshold is not reached then the registration is flagged as a failure. However, determining the value for the threshold would be difficult, particularly because the final values of pattern intensity and gradient difference not only depend on the registration accuracy, but are also effected by factors such as the size of the fluoroscopy ROI. If more than one registration is carried out, then it may be possible to use the previous value of the similarity measure to help calculate a threshold value.

It might also be possible to use multiple registrations to detect misregistrations. Starting from the assumption that local minima have a smaller capture range than the correct registration position. If a number of registrations are carried out, from different starting positions, then by analysing the spread of the final registration positions it may be possible to calculate a degree of confidence in the accuracy of a particular registration. Such a strategy would require a number of registrations and so the time taken to register would increase. However, Weese *et al.* [133] have achieved very rapid registrations (in 3 – 7 seconds) using shear warp factorisation [68, 69] and so a number of registrations could be carried out in a clinically useful time. Shear warp factorisation transforms the CT volume into a sheared object space by translating the CT slices relative to each other. The size of the translations depends on the viewing direction, the aim is to translate the slices so that the viewing direction is perpendicular to the slice plane. This enables the voxel slices to be projected in a more efficient manner. To account for perspective projection each image

slice is both translated and scaled.

As well as detecting misregistrations it is also important to detect whether the patient or the fluoroscopy set have moved, which would result in the current registration becoming invalid. Movement could be detected by continually calculating the similarity between a DRR produced at the registration position and the current fluoroscopy image. Apart from fluctuations due to noise, the value of the similarity measure should remain constant and a change in the similarity measure would indicate a misregistration. Similarly, the current fluoroscopy image could be compared to the fluoroscopy image on which the registration was based. Again, any large differences between the images would indicate a misregistration.

It may also be possible, once an initial registration has been achieved, to track the motion of the patient by carrying out successive registrations to each fluoroscopy frame, using the previous registration position as the starting position. A similar system has been demonstrated by Clarkson *et al.* [24] where a registration between video images and an MR image is tracked over a series of video frames.

Deformation

Deformation can be a problem when a pre-operative modality is used to represent an intra-operative scene. One method of countering this problem has been described in chapter 8, however, this work is in its preliminary stages. Out of the three interventional procedures described in chapter 2 the treatment of liver metastases is expected to have the largest problems caused by deformation. If patients being treated for liver metastases cannot reproducibly hold their breath to a sufficient precision, then a deformation algorithm would be required to enable an accurate registration. There are a number of possibilities for such an algorithm. It might be possible to use the algorithm described in chapter 8, where a number of 2D-3D registrations would be carried out on vertebrae and the ribs. The final extrinsic parameters would then be used to rigidly transform these bony structures allowing the soft tissue (e.g. the liver and metastases) to deform accordingly. Rueckert *et al.* [103] have designed a 3D-3D non-rigid registration algorithm using mutual information. An approach similar to Rueckert *et al.* [103] may be possible, where the CT scan is allowed to deform in order to further optimise the similarity between the DRRs and the fluoroscopy image. It may also be possible to employ a statistical shape model, such as the point distribution model [30, 31] to ascertain the most probably modes of deformation. This could be used to help constrain the search space during a non-rigid registration.

9.6 Overall Summary

The work described in this thesis has involved the design, implementation and validation of a 2D-3D registration algorithm for use in image guided interventions. The issues involved with the production of DRRs, the search space and search strategy have each been investigated and a number of novel contributions have been made. One of the major contributions of the work described in this thesis has been the comparison of registration accuracy and robustness when using different similarity measures. The two most accurate and robust measures were pattern intensity and gradient difference. Care has been taken to ensure that the algorithm could be directly used in a clinical setting. The starting positions used to validate the algorithm should be easy to acquire in most interventional procedures and the algorithm has been validated using routinely acquired clinical data. I believe that 2D-3D registration algorithms will become an important clinical tool in the future. These algorithms are most likely to be initially used in the field of radiotherapy, though there are a number of promising applications in surgical and interventional procedures. I hope and expect that either this algorithm, or some of the novel contributions which have gone into creating this algorithm, will enable more accurate clinical positioning in the future.

Appendix A

Accurately representing x-ray attenuation

This appendix describes how to account for the differences in fluoroscopy and CT x-ray energies by using the mixture rule. The aim is to find the relationship between the linear attenuation coefficient at the effective x-ray energy of a CT scanner (approximately 80keV) and the linear attenuation coefficient at the effective x-ray energy of a fluoroscopy set (approximately 50keV). Or, more specifically, starting with a Hounsfield number calculated using a CT scanner with an effective energy of about 80keV, H_{80}, the aim is to calculate what the Hounsfield number would be, if the same voxel was scanned with a CT scanner which used x-rays with an effective energy of about 50keV, H_{50}.

If an absorber is made up from a mixture of materials, then its mass attenuation coefficient μ/ρ can, to a very good approximation, be evaluated from the mass attenuation coefficients μ_i/ρ_i of the constituents of the mixture weighted according to their proportion by weight w_i. This is known as the mixture rule, see equation (A.1), where μ_i equals the linear attenuation coefficient and ρ_i represents the density of the i'th component of the mixture.

$$\mu = \rho \sum_i \frac{w_i \mu_i}{\rho_i} \tag{A.1}$$

Two assumptions are now made. Firstly, that the human body consists of only three components; air, water (soft tissue) and bone. Secondly, that only two of these three components can exist in a given voxel at one time. If the voxel has a Hounsfield value below zero it consists of water and air, and if the voxel has a Hounsfield value above zero it consists of water and bone. In both of these cases the volume of the voxel remains constant.

Consider the case of a voxel with a Hounsfield value below zero. Assuming that the fraction (by weight) of air w_{air} is negligible compared to w_{water} which is approximately unity, the linear attenuation coefficient of the voxel is given by equation (A.2). By substituting equation (A.2) into the equation to calculate Hounsfield values, equation (A.3) gives equation (A.4). This equation shows that the value of H is only dependent on the ratio of the density of the mixture to the density of water, therefore, for a voxel with a negative Hounsfield value the Hounsfield values are independent of x-ray energy, i.e. $H_{50} = H_{80}$.

$$\mu = \rho \frac{\mu_{water}}{\rho_{water}} \tag{A.2}$$

$$H = \frac{\mu - \mu_w}{\mu_w} 1000 \tag{A.3}$$

$$H = \left(\frac{\rho}{\rho_{water}} - 1 \right) 1000 \tag{A.4}$$

For a voxel which contains water and bone the linear attenuation coefficient is given by equation (A.5). Assuming a constant volume for each voxel V_{total} then, by substituting equation (A.5) into equation (A.3) it is possible to obtain equation (A.6), where V_{bone} is the volume of the voxel occupied by bone. This equation shows that for a given fraction of bone in a voxel the Hounsfield number does depend on the energy of the incident x-rays. It is a linear relationship and it is given by equation (A.7). Inputting values into this equation from Hubbell [62] results in a ratio of 1.67, therefore, positive Hounsfield numbers should be multiplied by 1.67 to account for the difference in x-ray energies.

$$\mu = \rho \left\{ w_{water} \frac{\mu_{water}}{\rho_{water}} + w_{bone} \frac{\mu_{bone}}{\rho_{bone}} \right\} \tag{A.5}$$

$$H = \frac{1000 V_{bone}}{V_{total} \mu_{water}} \left(\mu_{bone} - \mu_{water} \right) \tag{A.6}$$

$$\frac{H_{50}}{H_{80}} = \frac{\mu_{water80} (\mu_{bone50} - \mu_{water50})}{\mu_{water50} (\mu_{bone80} - \mu_{water80})} \tag{A.7}$$

If the effective energies of x-rays used by the two modalities are known accurately, then it should be possible to convert the Hounsfield numbers, from one effective energy to the other, reasonably accurately.

Bibliography

[1] A. Adam. The definition of interventional radiology (or, "when is a barium enema an interventional procedure?"). *Eur. Radiol.*, 8:1014–1015, 1998. Editorial.

[2] G. Alusi, A.C. Tan, A.D. Linney, K. Raoof, and A. Wright. Three dimensional tracking with ultrasound for augmented reality applications in skull base surgery. In *CVRMed-MRCAS '97*, pages 511–517, 1997.

[3] Z. Amin, S.G. Bown, and W.R. Lees. Local treatment of colorectal liver metastases: a comparison of interstital laser photocoagulation and percutaneous alcohol injection. *Clinical Radiology*, 48:166–171, 1993.

[4] S. Balter. An overview of digital radiography. In J.D. Newell, Jr and C.A. Kelsey, editors, *Digital imaging in diagnostic radiology*, pages 81–105. Church Livingstone Inc., 1990.

[5] R. Bansal, L. H. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. S. Duncan. A novel approach for the registration of 2D portal and 3D CT images for the treatment setup verification in radiotherapy. In *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '98*, pages 1075–1086, 1998.

[6] E. Barden. Quantitation of hepatic respiratory motion using ultrasound and magnetic resonance imaging. Intercalated BSc thesis, Division of Radiological Sciences, United Medical and Dental Schools of Guy's and St.Thomas's Hospitals, University of London., 1997.

[7] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[8] F. Betting and J. Feldmar. 3D-2D projective registration of anatomical surfaces with their projections. In *Proc. Information processing in medical imaging*, pages 275–286, 1995.

[9] J. Bijhold. Three-dimensional verification of patient placement during radiotherapy using portal images. *Med. Phys.*, 20(2):347–356, 1993.

[10] W. Birkfellner, F. Watzinger, F. Wanschitz, G. Enislidis, C. Kollmann, D. Rafolt, R. Nowotny, R. Ewers, and H. Bergmann. Systematic distortions in magnetic position digitizers. *Med. Phys.*, 25(11):2242–2248, 1998.

[11] R. Bissett, S. Boyko, K. Leszczynski, S. Cosby, P. Dunscombe, and N. Lightfoot. Radiotherapy portal verification: an observer study. *Br. J. Radiol.*, 68(806):165–174, 1995.

[12] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11:567–585, 1989.

[13] M. Börner, A. Lahmer, and U. Stier. Experiences with the robodoc system in more than 1000 cases. In *Computer Assisted Radiology and Surgery 1998*, pages 689–697, 1998.

[14] W.R. Bosch, D.A. Low, R.L. Gerber, J.M. Michalski, M.V. Graham, C.A. Perez, W.B. Harms, and J.A. Purdy. The electronic view box: a software tool for radiation therapy treatment verification. *Int. J. Radiat. Oncol. Biol. Phys.*, 31(1):135–142, 1995.

[15] C.B. Bose and I. Amir. Design of fiducials for accurate registration using machine vision. *IEEE Trans. Pattern Anal. Mach. Intell*, 12:1196–1200, 1990.

[16] M. Breeuwer, J. P. Wadley, H. L. T. de Bliek, P. A. C. Desmedt, F. A. Gerritsen, P. Gieles, J. Buurman, N. L. Dorward, B. Velani, N. D. Kitchen, D. G. T. Thomas, O. Wink, J. D. Blankensteijn, B. C. Eikelboom, W. P. Th. M. Mali, M. A. Viergever, G. P. Penney, R. Gaston, C. R. Maurer, Jr., D. L. G. Hill, D. J. Hawkes, F. Maes, D. Vandermeulen, R. Verbeeck, P. Suetens, J. Weese, J. Sabczynski, W. Zylka, C. Lorenz, T. M. Buzug, G. Schmitz, and M. H. Kuhn. The EASI Project—Improving the effectiveness and quality of image-guided surgery. *IEEE Trans. Inform. Technol. Biomedicine*, 2:156–168, 1998.

[17] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24:325–375, 1992.

[18] T.M. Buzug, J. Weese, C. Fassnacht, and C. Lorenz. Image registration: Convex weighting functions for histogram-based similarity measures. In J. Troccaz, E. Grimson, and R. Mösges, editors, *Proc. CVRMed/MRCAS*, pages 203–212. Berlin, Germany: Springer-Verlag, 1997.

[19] G. Champleboux, S. Lavallée, P. Sautot, and P. Cinquin. Accurate calibration of cameras and range imaging sensors, the NPBS method. In *IEEE international conference on robotics and automation*, pages 1552–1558, May 1992.

[20] F. Chassat and Lavallée. Experimental protocol for accuracy evaluation of 6–D localizers for computer–integrated surgery: application to four optical localizers. In *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '98*, pages 277–284. Springer, 1998.

[21] C.W. Cheng, L.M. Chin, and P.K. Kijewski. A coordinate transfer of anatomical information from CT to treatment simulation. *Int. J. Radiat. Oncol. Biol. Phys.*, 13:1559–1569, 1987.

[22] G. Chiorboli and G.P. Vecchi. Comments on "design of fiducials for accurate registration using machine vision". *IEEE Trans. Pattern Anal. Mach. Intell*, 15:1330–1332, 1993.

[23] D.S.J. Choy, P.W. Ascher, S. Saddekni, D. Alkaitis, W. Liebler, J. Hughes, S. Diwan, and P. Altman. Percutaneous laser disc decompression. *Spine*, 17:949–956, 1992.

[24] M.J. Clarkson, D. Rueckert, A.P. King, P.J. Edwards, D.L.G. Hill, and D.J. Hawkes. Using texture mapping to register video images to tomographic images by optimising mutual information. In *Proc. Medical Image Understanding and Analysis '99*, pages 29–32, 1999.

[25] A.C.F. Colchester, J. Zhao, C. Henri, R.L. Evans, P. Roberts, N. Maitland, D.J. Hawkes, D.L.G. Hill, A.J. Strong, D.G. Thomas, M.J. Gleeson, and T.C.S. Cox. Craniotomy simulation and guidance using a stereo based tracking system (vislan). In *Proc. SPIE*, volume 2359, pages 541–551, 1994.

[26] A.C.F. Colchester, J. Zhao, K.S. Holton-Tainter, C. Henri, N. Maitland, P.T.E. Roberts, C.G. Harris, and R.J. Evans. Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Med. Image Anal.*, 1(1):73–90, 1996.

[27] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Information Processing in Medical Imaging 1995*, pages 263–274, 1995.

[28] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3D multi-modality medical image registration using feature space clustering. In *Computer Vision, Virtual Reality, and Robotics in Medicine 1995*, pages 195–204, 1995.

[29] D.L. Collins, T.M. Peters, and A.C. Evans. An automated 3D non-linear deformation procedure for determination of gross morphometric variability in human brain. In *Visualization in Biomedical Computing 1994*, pages 180 – 190, 1994.

[30] T.F. Cootes, D.H. Cooper, C.J. Taylor, and J. Graham. A trainable method of parametric shape description. *Image and Vision Computing*, 10:289–294, 1992.

[31] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. BMVC., Leeds, Springer Verlag*, pages 9–18, 1992.

[32] M.G. Cowling, N.C. Cheung, T.J. Gibson, and A.N. Adam. Percutaneous laser nucleotomy – techniques and preliminary experience. In *Proc. RSNA 83rd scientific assembly and annual meeting, Chicago, Illinois.*, volume 205(P), page 405, 1997. Abstract, 1116.

[33] D.R. Dance. Diagnostic radiology with x-rays. In S. Webb, editor, *The physics of medical imaging*, chapter 2. IOP publishing Ltd., 1988.

[34] J.T. De Sanctis, S.N. Goldberg, and P.R. Müller. Percutaneous treatment of hepatic neoplasms: a review of current techniques. *Cardiovascular and interventional radiology*, 21:273–296, 1998.

[35] F.C. Decibel and D.J. Monyak. Simplified transfer of anatomical information from CT scan to simulator film. *Med. Phys.*, 14(3):485, 1987. Meeting Abstract.

[36] J. Dimnet, L.P. Fischer, G. Gonon, and J.P. Carret. Radiographic studies of lateral flexion in the lumbar spine. *J. Biomechanics*, 11:143–150, 1978.

[37] L. Dong and A.L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1053–1060, 1995.

[38] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. Wiley-Interscience, 1973.

[39] P.J. Edwards, A.P. King, C.R. Maurer, Jr., D.A. de Cunha, M.R. Pike, D.J. Hawkes, D.L.G. Hill, R.P. Gaston, M.R. Fenlon, A.J. Strong, C.L. Chandler, and M.J. Gleeson. A locking acrylic dental stent for registration and tracking during image-guided surgery of the head. In *Proc. Medical Image Understanding and Analysis '99*, pages 105–108, 1999.

[40] R. Fahrig, M. Moreau, and D.W. Holdsworth. Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: Correction of image intensifier distortion. *Med. Phys.*, 24(7):1097–1106, 1997.

[41] R.F. Farr and P.J. Allisy-Roberts. *Physics for Medical Imaging*. WB Saunders, 1998.

[42] J. Feldmar, N. Ayache, and F. Betting. 3D-2D projective registration of free-form curves and surfaces. *Comput. Vision Image Understanding*, 65(3):403–424, 1997.

[43] J. Feldmar, G. Malandain, N. Ayache, S. Fernández-Vidal, E. Maurincomme, and Y. Trousset. Matching 3D MR angiography data and 2D x-ray angiograms. In *CVRMed-MRCAS '97*, pages 129–138, 1997.

[44] S. Ganapathy. Decomposition of transformation matrices for robot vision. *Pattern Recognition Letters 2*, pages 401–412, 1984.

[45] K. G. A. Gilhuijs, P. J. H. van de Ven, and M. van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Med. Phys.*, 23(3):389–399, 1996.

[46] K.G.A. Gilhuijs and M. van Herk. Automatic on-line inspection of patient setup in radiation therapy using digital portal images. *Med. Phys.*, 20(3):389–399, 1993.

[47] G.H. Golub and C.F. Van Loan. *Matrix computations*. The John Hopkins University Press, third edition, 1996.

[48] R.C. Gonzalez and R.E. Woods. *Digital Image Processing*. Addison Wesley, 1992.

[49] L.M. Gottesfeld Brown and T.E. Boult. Registration of planar film radiographs with computed tomography. In *Proc. MMBIA*, pages 42–51, 1996.

[50] D. Grönemeyer, A. Melzer, and R. Seibel. Interventional CT and MRI combined with microendoscopy for low back pain therapy. *Min. Invas. Ther. & Allied Technol.*, 7(3):223–232, 1998.

[51] A. Guéziec, P Kazanzides, B. Williamson, and R. H. Taylor. Anatomy based registration of CT-scan and x-ray images for guiding a surgical robot. *IEEE Trans. Med. Imaging*, 17(5):715–728, 1998.

[52] A. Guéziec, P Kazanzides, B. Williamson, R. H. Taylor, and D. Lord. Anatomy based registration of CT-scan and x-ray fluoroscopy data for intra-operative guidance of a surgical robot. In *Proc. SPIE Medical Imaging*, pages 81–94, 1998.

[53] P. Haaker, E. Klotz, R. Koppe, and R. Linde. Real-time distortion correction of digital x-ray II/TV-systems: an application example for digital flashing tomosynthesis (DFTS). *International Journal of cardiac imaging*, 6:39–45, 1990.

[54] J.V. Hajnal, N. Saeed, E.J. Soar, A. Oatridge, I.R. Young, and G.M. Bydder. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *J. Comput. Assist. Tomogr.*, 19(2):289–296, 1995.

[55] A. Hamadeh, P. Sautot, and P. Cinquin. A unified approach to 3D-2D registration and 2D images segmentation. In *Computer Assisted Radiology 1995*, pages 1191–1196, 1995.

[56] S. B. Harvey, F. W. Smith, and D. W. L. Hukins. Measurement of lumbar spine flexion-extension using a low-field open-magnet magnetic resonance scanner. *Investigative radiology*, 33(8):439–443, 1998.

[57] D. L. G. Hill, D. J. Hawkes, Z. Hussain, S.E.M. Green, C.F. Ruff, and G.P. Robinson. Accurate combination of CT and MR data of the head: validation and applications in surgical and therapy planning. *Computerised Medical Imaging and Graphics.*, 17(4/5):357–363, 1993.

[58] D. L. G. Hill, C. R. Maurer, Jr., R. J. Maciunas, J. A. Barwise, J. M. Fitzpatrick, and M. Y. Wang. Measurement of intraoperative brain surface deformation under a craniotomy. *Neurosurgery*, 43(3):514–528, 1998.

[59] D. L. G. Hill, C. R. Maurer, Jr., C. Studholme, J. M. Fitzpatrick, and D. J. Hawkes. Correcting scaling errors in tomographic images using a nine degree of freedom registration algorithm. *J. Comput. Assist. Tomogr.*, 22(2):317–323, 1998.

[60] G.N. Hounsfield. Computerized transverse axial scanning (tomography): Part 1. description of system. *Br. J. Radiol.*, 46:1016–1022, 1973.

[61] D.H. Hristov and B.G. Fallone. A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Med. Phys.*, 23(1):75–84, 1996.

[62] J.H Hubbell. *Photon Cross Sections, Attenuation Coefficients, and Energy Absorption Coefficients From 10keV to 100GeV*. Center for Radiation Research, National Bureau of Standards, Washington, D.C.20234, 29 edition, 1969.

[63] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *Proc. International Conference on Computer Vision*, pages 259–268, 1987.

[64] P.J. Kelly and B.A. Kall, editors. *Computers in Stereotactic Neurosurgery*. Blackwell Scientific Publications, 1992.

[65] Y. Kita, D. L. Wilson, and J. A. Noble. Real-time registration of 3D cerebral vessels to x-ray angiograms. In *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '98*, pages 1125–1133, 1998.

[66] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proc. International Joint Conference on Artificial Intelligence*, pages 1137–1143, 1995.

[67] R.A. Kruger, C.A. Mistretta, and S.J. Riederer. Physical and technical considerations of computerized fluoroscopy difference imaging. *IEEE Trans. Nucl. Sci.*, NS-28:205–212, 1981.

[68] P. Lacroute and M. Levoy. Fast volume rendering using a shear-warp factorization of the viewing transformation. In *Proc. SIGGRAPH '94*, pages 451–458, 1994. Computer Graphics Proceedings, ACM SIGGRAPH, New York.

[69] P.G. Lacroute. *Fast Volume Rendering Using a Shear Warp Factorization of theViewing Transformation*. PhD thesis, Stanford University, 1995.

[70] S. Lavallée and R. Szeliski. Recovering the position and orientation of free-form objects from image contours using 3-D distance maps. *IEEE Trans. Pattern Anal. Mach. Intell*, 17(4):378–390, 1995.

[71] S. Lavallée, R. Szeliski, and L. Brunie. Matching 3-D smooth surfaces with their 2-D projections using 3-D distance maps. In *Proc. SPIE Geometric methods in computer vision.*, volume 1570, pages 322–336, 1991.

[72] L. Lemieux and R. Jagoe. Effect of fiducial marker localization on stereotactic target coordination in CT slices and radiographs. *Phys. Med. Biol.*, 39:1915–1928, 1994.

[73] L. Lemieux, R. Jagoe, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med. Phys.*, 21(11):1749–1760, 1994.

[74] W.A. Liebler. Percutaneous laser disc nucleotomy. *Clinical Orthopaedics and Related Research*, (310):58–66, 1995.

[75] J. A. Little, D. L. G. Hill, and D. J. Hawkes. Deformations incorporating rigid structures. *Computer Vision and Image Understanding*, 66(2):223–232, 1997.

[76] J.A. Little and D.J. Hawkes. The registration of multiple medical images acquired from a single subject: why, how, what next? *Statistical Methods in Medical Research*, 6:239–265, 1997.

[77] A. Liu, E. Bullitt, and S. M. Pizer. Surgical instrument guidance using synthesized anatomical structures. In *Proc. CVRMed/MRCAS*, pages 99–108, 1997.

[78] A. Liu, E. Bullitt, and S. M. Pizer. 3D/2D registration via skeletal near projective invariance in tubular objects. In *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '98*, pages 952–963, 1998.

[79] A. Liu, S.M. Pizer, D. Eberly, B. Morse, J. Rosenman, E. Chaney, E. Bullitt, and V. Carrasco. Volume registration using the 3D core. In *Proc. Visualization in Biomedical Computing*, pages 217–226, 1994.

[80] A.E. Lujan, J.M. Balter, and R.K. Ten Haken. Determination of rotations in three dimensions using two-dimensional portal image registration. *Med. Phys.*, 25(5):703–708, 1998.

[81] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Seutens. Multimodality image registration by maximization of mutual, information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[82] M. Makaroun, P. Zajko, A. Orons, S. Muluk, R. Rhee, D. Steed, and M. Webster. The experience of an academic medical center with endovascular treatment of abdominal aortic aneurysms. *Am. J. Surgery*, 176:198–202, August 1998.

[83] K. V. Mardia, J. T. Kent, C. R. Goodall, and J. A. Little. Kriging and splines with derivative information. *Biometrika*, 83(1):207–221, 1996.

[84] S.R. Marschner and R.J. Lobb. An evaluation of reconstruction filters for volume rendering. In R.D. Bergeron and A.E. Kaufman, editors, *Proc. Visualization 94*, pages 100–107. IEEE Computer Society Press, 1994.

[85] Y. Masutani, T. Dohi, F. Yamane, H. Iseki, and K. Takakura. Interactive virtualized display system for intravascular neurosurgery. In *CVRMed-MRCAS '97*, pages 427–435, 1997.

[86] C. R. Maurer, Jr., D. L. G. Hill, A. J. Martin, H. Liu, M. McCue, D. Rueckert, D. Lloret, W. A. Hall, R. E. Maxwell, D. J. Hawkes, and C. L. Truwit. Investigation of intraoperative brain deformation using a 1.5 Tesla interventional MR system: Preliminary results. *IEEE Trans. Med. Imaging*, 17(5):817–825, 1998.

[87] C.R. Maurer, Jr., G.B. Aboutanos, B.M. Dawant, Maciunas R.J., and J.M. Fitzpatrick. Registration of 3-D images using weighted geometrical features. *IEEE Trans. Med. Imaging*, 15(6):836–849, 1996.

[88] C.R. Maurer, Jr. and J.M. Fitzpatrick. *Interactive Image-Guided Neurosurgery*, chapter A review of medical image registration, pages 17–44. American Association of Neurological Surgeons, Park Ridge, IL, 1993.

[89] C.R. Maurer, Jr., J.M. Fitzpatrick, M.Y. Wang, R.L. Galloway, Jr., Maciunas R.J., and G.S. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Trans. Med. Imaging*, 16:447–462, 1997.

[90] J. May, G. White, R. Waugh, W. Yu, and J. Harris. Treatment of complex abdominal aortic aneurysms by a combination of endoluminal and extraluminal grafts. *J. of Vascular Surgery*, 19(5):924–933, 1995.

[91] C. A. Micchelli. Interpolation of scattered data: Distance matrices and conditionally positive definite functions. *Constructive Approximation*, 2:11–22, 1986.

[92] D.F. Morrison. *Multivariate statistical methods*. McGraw-Hill International Editions, third edition, 1990.

[93] K.D. Murphy, G.M. Richter, M. Henry, C.E. Encarnacion, V.A. Le, and J.C. Palmaz. Aortoiliac aneurysms : management with endovascular stent-graft placement. *Radiology*, 198:437–480, 1996.

[94] M. J. Murphy. An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery. *Med. Phys.*, 24(6):857–866, June 1997.

[95] D. E. Myers. Kriging, cokriging, radial basis functions and the role of positive definiteness. *Computers Math. Applic.*, 24(12):139–148, 1992.

[96] J.A. Parker, R.V. Kenyon, and D.E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imaging*, M1-2(1):31–39, 1983.

[97] G. P. Penney, J. A. Little, R. P. Gaston, M. G. Cowling, D. L. G. Hill, and D. J. Hawkes. The registration of images of the spine after patient movement using deformations incorporating rigid structures: preliminary results. In *Proc. Medical Image Understanding and Analysis '98*, pages 41–44, 1998. http://www.robots.ox.ac.uk/ mvl/miua98/abstract41.html.

[98] T.Q. Phong, R. Horaud, A. Yassine, and P.D. Tao. Object pose from 2-D to 3-D point and line correspondences. *Int. J. Comput. Vision*, 15:225–243, 1995.

[99] M.S. Potsaid. *Interventional Radiology*, chapter 61, pages 785–789. W.B. Saunders Company, 1982.

[100] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, 1992.

[101] T. Radcliffe, R. Rajapakshe, and S. Shalev. Pseudocorrelation: a fast, robust, absolute, grey-level image alignment algorithm. *Med. Phys.*, 21(6):761–769, 1994.

[102] A. Rougée, C. Picard, C. Ponchut, and Y. Trousset. Geometric calibration of X-ray imaging chains for three-dimensional reconstruction. *Computerized Medical Imaging and Graphics*, 17(4/5):295–300, 1993.

[103] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D.J. Hawkes. Non-rigid registration of breast MR images using mutual information. In *Medical Imaging Computing and Computer-Assisted Intervention—MICCAI '98*, pages 1144–1152, 1998.

[104] D. Ruprecht, R. Nagel, and H. Müller. Spatial free-form deformation with scattered data interpolation methods. *Computers and Graphics*, 19(1):63–71, 1995.

[105] D.R. Sandeman, N. Patel, C. Chandler, R.J. Nelson, H.B. Coakham, and H.B. Griffith. Advances in image-directed neurosurgery - preliminary experience with the ISG viewing wand compared with the Leksell-G frame. *Br. J. Neurosurgery*, 8(5):529–544, 1994.

[106] S. Schreiner, J. Funda, A.C. Barnes, and J.H. Anderson. Accuracy assessment of a clinical biplane fluoroscope for three-dimensional measurements and targeting. In *Medical Imaging 1997: Image Display*, pages 160–166, 1997.

[107] D. Shepard. A two-dimensional interpolation function for irregularly spaced data. In *Proc. 23rd Nat'l Conf. of the ACM*, pages 517–524, New York, 1968. ACM Press.

[108] G.W. Sherouse, K. Novins, and E.L. Chaney. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int. J. Radiation Oncology Biol. Phys.*, 18:651–658, 1990.

[109] C.L. Siegel and R.H. Cohan. CT of abdominal aortic aneurysms. *American Journal of Roentgenology*, 163(1):17–29, 1994.

[110] L. Solbiati, T. Ierace, S.N. Goldberg, et al. Percutaneous US-guided radio-frequency tissue ablation of liver metastases: Treatment and follow-up in 16 patients. *Radiology*, 202:195–203, 1997.

[111] T.M. Strat. Recovering the camera parameters from a transformation matrix. In *Proc. DARPA image understanding workshop*, pages 264–271, 1984.

[112] C. Studholme. *Measures of 3D Medical Image Alignment*. PhD thesis, United Medical and Dental Schools of Guy's and St. Thomas's Hospitals, 1997.

[113] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3D registration of truncated MR and CT images of the head. In D. Pycock, editor, *Proc. Br. Mach. Vision Conf.*, volume 1, pages 27–37, 1995.

[114] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.

[115] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med. Phys.*, 24(1):25–35, 1997.

[116] C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.

[117] C. Studholme, J.A. Little, G.P. Penney, D.L.G. Hill, and D.J. Hawkes. Automated multi-modality registration using the full affine transformation: application to MR and CT guided skull based surgery. In *Visualization in Biomedical Computing 1996*, pages 601–606, 1996.

[118] R. H. Taylor, S. Lavallée, G. C. Burdea, and R. W. Mosges, editors. *Registration for Computer Integrated Surgery:Methodology, state of the art*, pages 425–449. Cambridge, MA : MIT Press, 1995.

[119] R.H. Taylor, L. Joskowicz, B. Williamson, A. Guéziec, A. Kalvin, P. Kazanaides, R. Van Vorhis, J. Yao, R. Kumar, A. Bzostek, A. Sahay, M. Börner, and A. Lahmer. Computer-integrated revision total hip replacement surgery: concept and preliminary results. *Med. Image Anal.*, 3(3):301–319, 1999.

[120] D.G.T. Thomas, S.S. Gill, C.B. Wilson, J.L. Darling, and C.S. Parkins. Use of relocatable stereotactic frame to integrate positron emission tomography and computed tomography

images: application in human malignant brain tumours. In *Proc. Stereotactic Funct. Neurosurgery*, pages 388–392, 1990.

[121] G.S. Treiman and V.M. Bernhard. Endovascular treatment of abdominal aortic aneurysms. *Annu. Rev. Med.*, 49:363–73, 1998.

[122] R.Y. Tsai. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off the shelf TV cameras and lenses. *IEEE Trans. Robotics Automat.*, RA-3(4):323–344, 1987.

[123] P.A. van den Elsen, E.J.D. Pol, and M.A. Viergever. Medical image matching — a review with classification. *IEEE Eng. Med. Biol.*, 12(1):26–39, March 1993.

[124] P. Viola and W. M. Wells. Alignment by maximization of mutual information. In *Proc. 5th International Conference on Computer Vision (ICCV'95)*, pages 16–23, Boston, U.S.A., 1995.

[125] T.J. Vogl, P.K. Muller, R. Hammerstimgl, N. Weinhold, M.G. Mack, C. Philipp, M. Deimling, J. Beuthan, W. Pegios, H. Riess, H. Lemmens, and R. Felix. Malignant liver tumours treated with MR image guided laser-induced thermotherapy: techniques and prospective results. *Radiology*, 196(1):257–265, 1995.

[126] T.L. Vogl, M.G. Mack, R. Straub, A. Roggan, and R. Felix. Magnetic resonance imaging — guided abdominal interventional radiology: laser–induced thermotherapy of liver metastases. *Endoscopy*, 29:577–583, 1997.

[127] F. Vorbeck, M. Cartellieri, K. Ehrenberger, and H. Imhof. Experiences in intraoperative computer-aided navigation in ENT sinus surgery with the aesculap navigation system. *Comput. Aided Surg.*, 3:306–311, 1998.

[128] J.S. Wagner, M.A. Adson, J.A. Van Heerden, M.H. Adson, and D.M. Ilstrup. The natural history of hepatic metastases from colorectal cancer. A comparison with resective treatment. *Annals of Surgery*, 199:502–508, 1984.

[129] J.Z. Wang, L.E. Reinstein, J. Hanley, and A.G. Meek. Investigation of a phase-only correlation technique for anatomical alignment of portal images in radiation therapy. *Phys. Med. Biol.*, 41:1045–1058, 1996.

[130] M.Y. Wang, C.R. Maurer, Jr., and J.M. Fitzpatrick. Partial volume effect on marker localisation in medical density images. In *SPIE*, volume 3034, pages 580–591, 1997.

[131] J. Weese, T. M. Buzug, C. Lorenz, and C. Fassnacht. An approach to 2D/3D registration of a vertebra in 2D x-ray fluoroscopies with 3D CT images. In *Proc. CVRMed/MRCAS*, pages 119–128, 1997.

[132] J. Weese, T.M. Buzug, G.P. Penney, and D.J. Hawkes. Method for determining the transformation between an object and its three dimensional representation and device for carrying out this method. Patent extension JP10323347. Extension to European Patent DE19705600.

[133] J. Weese, R. Göcke, G.P. Penney, P. Desmedt, T.M. Buzug, and H. Schumann. Fast voxel-based 2D/3D registration using a volume rendering method based on shear-warp factorization. In *Medical Imaging 1999: Image Processing*, pages 802–810. SPIE–The international society for optical engineering, 1999.

[134] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes. Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. *IEEE Trans. on Information Technology in Biomedicine*, 1(4):284–293, 1997.

[135] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(10):965–980, 1992.

[136] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer, Jr., R.M. Kessler, and R.J. Maciunas. Retrospective intermodality registration techniques for images of the head: surface-based verses volume-based. *IEEE Trans. Med. Imaging*, 18(2):144–150, 1999.

[137] J. B. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods. Comparison and evaluation of retrospective intermodality image registration techniques. *J. Comput. Assist. Tomogr.*, 21:554–566, 1997.

Publications

Articles in journals

- M. Breeuwer, J.P. Wadley, H.L.T. de Bliek, P.A.C. Desmedt, F.A. Gerritsen, P. Gieles, J. Buurman, N.L. Dorward, B. Velani, N.D. Kitchen, D.G.T. Thomas, O. Wink, J.D. Blankensteijn, B.C. Eikelboom, W.P Th. M Mali, M.A. Viergever, G.P. Penney, R. Gaston, C.R. Maurer Jr., D.L.G. Hill, D.J. Hawkes, F. Maes, D. Vandermeulen, R. Verbeeck, P. Suetens, J. Weese, J. Sabczynski, W. Zylka, C. Lorenz, T.M. Buzug, G. Schmitz and M.H. Kuhn. The EASI Project—Improving the effectiveness and quality of image-guided surgery, *IEEE Trans. Inform. Technol. Biomedicine* and *Proc. Computer Assisted Radiology and Surgery 1998*, 2(3):156–168,1998

- J.Weese, T.M.Buzug, G.P.Penney and P.Desmedt. 2D/3D registration and motion tracking for surgical interventions, *Philips Journal of Research*, 51(2):299–316, 1998.

- G.P.Penney, J.Weese, J.A.Little, P.Desmedt, D.L.G.Hill and D.J.Hawkes. A comparison of similarity measures for use in 2D-3D medical image registration. *IEEE Trans. Med. Imaging* and *Proc. Medical Image Computing and Computer Assisted Intervention 1998*, 17(4):586–595, 1998.

- J.Weese, G.P.Penney, P.Desmedt, T.M.Buzug, D.L.G.Hill and D.J.Hawkes. Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for inage-guided surgery. *IEEE Trans. Inform. Technol. Biomedicine*, 1(4):284–293, 1997.

Articles in conference procedings

- G.P.Penney, J.A.Little, J.Weese, D.L.G.Hill and D.J.Hawkes. Deforming a pre-operative volume to represent the intra-operative scene. *Proc. SPIE Medical Imaging 2000: Image Processing*.

- J.Weese, R.Göcke, G.P.Penney, P.Desmedt and H.Schumann. Fast voxel based 2D/3D registration using a volume rendering method based on shear-warp factorization. *Proc. SPIE Medical Imaging 1999: Image Processing*, pages 802–810, 1999.

- G.P.Penney, J.A.Little, R.G.Gaston, M.G.Cowling, D.L.G.Hill and D.J.Hawkes. The registration of images of the spine after patient movement using deformations incorporating rigid structures: preliminary results. In *Proc. Medical Image Understanding and Analysis '98*, pages 41–44, 1998.

- J.Weese, G.P.Penney, T.M.Buzug, C.Fassnacht and C.Lorentz. 2D/3D registration of pre-operative CT images and intra-operative x-ray projections for image guided surgery. in *Proc. Computer assisted Radiology and Surgery*, pager 833–838, 1997.

- G.P.Penney, J.A.Little, D.L.G.Hill and D.J.Hawkes. 2D-3D image registration for use in image guided interventions. In *Proc. Medical Image Understanding and Analysis '97*, pages 49–52, 1997.

- C.Studholme, J.A.Little G.P.Penney, D.L.G.Hill and D.J.Hawkes. Automated multimodality registration using the full affine transformation: application to MR and CT guided skull based surgery. In *Proc. Visualization in Biomedical Computing*, pages 601–606, 1996.

Patent application

- J.Weese, T.M.Buzug, G.P.Penney and D.J.Hawkes. Method for determining the transformation between an object and its three dimensional representation and device for carrying out this method. Patent extension JP10323347. Extension to European Patent DE19705600.